GenCore version 5.1.4_p5_4578

OM protein - protein search, using sw model

Run on: April 2, 2003, 14:48:15 ; Search time 28 Seconds
(without alignments)
270.060 Million cell updates/sec

Title: US-10-002-278-9
Perfect score: 1368
Sequence: 1 DVLEDSETWDQATGTKTFLV.....GVPTLKYHYEGMSVAECGCR 257

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA:*
1: /cgn2_6/ptodata/1/iaa/5A_COMB.pep:*
2: /cgn2_6/ptodata/1/iaa/5B_COMB.pep:*
3: /cgn2_6/ptodata/1/iaa/6A_COMB.pep:*
4: /cgn2_6/ptodata/1/iaa/6B_COMB.pep:*
5: /cgn2_6/ptodata/1/iaa/PCTUS_COMB.pep:*
6: /cgn2_6/ptodata/1/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	1368	100.0	257	4	US-08-065-844A-9	Sequence 9, Appli
2	1336.5	97.7	428	1	US-08-050-132A-2	Sequence 2, Appli
3	1336.5	97.7	428	3	US-08-750-222A-2	Sequence 2, Appli
4	1336.5	97.7	428	3	US-08-815-652B-2	Sequence 2, Appli
5	1336.5	97.7	428	4	US-08-254-353A-2	Sequence 2, Appli
6	1336.5	97.7	428	5	PCT-US92-05374A-2	Sequence 2, Appli
7	1336.5	97.7	428	5	PCT-US95-07084-2	Sequence 2, Appli
8	808.5	59.1	427	4	US-08-065-844A-2	Sequence 2, Appli
9	634	46.3	150	1	US-08-050-132A-9	Sequence 9, Appli
10	634	46.3	150	3	US-08-750-222A-9	Sequence 9, Appli
11	634	46.3	150	3	US-08-815-652B-9	Sequence 9, Appli
12	634	46.3	150	4	US-08-254-353A-9	Sequence 9, Appli
13	634	46.3	150	5	PCT-US92-05374A-9	Sequence 9, Appli
14	634	46.3	150	5	PCT-US95-07084-9	Sequence 9, Appli
15	519.5	38.0	312	1	US-08-247-908A-2	Sequence 2, Appli
16	519.5	38.0	312	1	US-08-453-942-2	Sequence 2, Appli
17	519.5	38.0	312	2	US-08-926-885A-2	Sequence 2, Appli
18	519.5	38.0	312	5	PCT-US94-05290-2	Sequence 2, Appli
19	516.5	37.8	424	1	US-08-247-908A-11	Sequence 11, Appl
20	516.5	37.8	424	1	US-08-453-942-11	Sequence 11, Appl
21	516.5	37.8	424	2	US-08-926-885A-11	Sequence 11, Appl
22	516.5	37.8	424	5	PCT-US94-05290-11	Sequence 11, Appl
23	469	34.3	103	1	US-08-335-583C-53	Sequence 53, Appl
24	469	34.3	103	3	US-08-478-097A-15	Sequence 15, Appl
25	469	34.3	103	4	US-08-931-858E-168	Sequence 168, App
26	469	34.3	103	4	US-08-981-739-168	Sequence 168, App
27	469	34.3	103	4	US-09-128-026-168	Sequence 168, App
28	349.5	25.5	321	1	US-08-362-670B-26	Sequence 26, Appl
29	349.5	25.5	321	3	US-08-333-576C-26	Sequence 26, Appl
30	349.5	25.5	321	4	US-08-808-324-26	Sequence 26, Appl
31	349.5	25.5	321	5	PCT-US94-14030A-26	Sequence 26, Appl
32	348.5	25.5	417	1	US-08-199-780-1	Sequence 1, Appli
33	348.5	25.5	417	2	US-08-316-650-1	Sequence 1, Appli
34	348.5	25.5	417	5	PCT-US95-02251-1	Sequence 1, Appli
35	345	25.2	408	1	US-07-841-646-7	Sequence 7, Appli
36	345	25.2	408	1	US-07-718-274A-6	Sequence 6, Appli
37	345	25.2	408	1	US-07-901-703-19	Sequence 19, Appl
38	345	25.2	408	1	US-08-149-106-6	Sequence 6, Appli
39	345	25.2	408	1	US-08-147-023-7	Sequence 7, Appli
40	345	25.2	408	1	US-08-298-021-6	Sequence 6, Appli
41	345	25.2	408	1	US-08-050-132A-4	Sequence 4, Appli
42	345	25.2	408	1	US-08-447-570-7	Sequence 7, Appli
43	345	25.2	408	2	US-08-459-346-11	Sequence 11, Appl
44	345	25.2	408	2	US-08-449-700-7	Sequence 7, Appli
45	345	25.2	408	2	US-07-989-847-4	Sequence 4, Appli

ALIGNMENTS

```
RESULT 1
US-08-065-844A-9
; Sequence 9, Application US/08065844A
; Patent No. 6333168
;  GENERAL INFORMATION:
;    APPLICANT:  Jessell, Thomas M.
;    APPLICANT:  Basler, Konrad
;    APPLICANT:  Yomada, Toshiya
;    TITLE OF INVENTION:  CLONING, EXPRESSION AND USES OF
;    TITLE OF INVENTION:  DORSALIN-1
;    NUMBER OF SEQUENCES:  18
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Cooper & Dunham
;      STREET:  30 Rockefeller Plaza
;      CITY:  New York
;      STATE:  New York
;      COUNTRY:  United States of America
;      ZIP:  10112
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/065,844A
;      FILING DATE:  19930520
;      CLASSIFICATION:  424
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  White, John P.
;      REGISTRATION NUMBER:  28,678
;      REFERENCE/DOCKET NUMBER:  0576/40314
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (212) 977-9550
;      TELEFAX:  (212) 664-0525
;      TELEX:  422523 COOP UI
;  INFORMATION FOR SEQ ID NO:  9:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  257 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
;    HYPOTHETICAL:  NO
;    ANTI-SENSE:  NO
US-08-065-844A-9

  Query Match           100.0%;  Score 1368;  DB 4;  Length 257;
  Best Local Similarity  100.0%;  Pred. No. 1.5e-131;
  Matches  257;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;
```

```
Qy      1 DVLEDSETWDQATGTKTFLVSQDIRDEGWETLEVSSAVKRWVRADSTTNKNKLEVTVQSH 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 DVLEDSETWDQATGTKTFLVSQDIRDEGWETLEVSSAVKRWVRADSTTNKNKLEVTVQSH 60

Qy     61 RESCDTLDISVPPGSKNLPFFVVFSNDRSNGTKETRLDLLKEMIGHEQETMLVKTAKNAY 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 RESCDTLDISVPPGSKNLPFFVVFSNDRSNGTKETRLDLLKEMIGHEQETMLVKTAKNAY 120

Qy    121 QGAGESQEEEGLDGYTAVGPLLARRKRSTGASSHCQKTSLRVNFEDIGWDSWIIAPKEYD 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 QGAGESQEEEGLDGYTAVGPLLARRKRSTGASSHCQKTSLRVNFEDIGWDSWIIAPKEYD 180

Qy    181 AYECKGGCFFPLADDVTPTKHAIVQTLVHLKFPTKVGKACCVPTKLSPISILYKDDMGVP 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 AYECKGGCFFPLADDVTPTKHAIVQTLVHLKFPTKVGKACCVPTKLSPISILYKDDMGVP 240

Qy    241 TLKYHYEGMSVAECGCR 257
          |||||||||||||||||
Db    241 TLKYHYEGMSVAECGCR 257


RESULT 2
US-08-050-132A-2
; Sequence 2, Application US/08050132A
; Patent No. 5661007
;  GENERAL INFORMATION:
;    APPLICANT:  Wozney, John M.
;    APPLICANT:  Celeste, Anthony
;    TITLE OF INVENTION:  BMP-9 COMPOSITIONS
;    NUMBER OF SEQUENCES:  19
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Genetics Institute, Inc.
;      STREET:  Legal Affairs - 87 CambridgePark Drive
;      CITY:  Cambridge
;      STATE:  MA
;      COUNTRY:  US
;      ZIP:  02140
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/050,132A
;      FILING DATE:
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Kapinos, Ellen J.
;      REGISTRATION NUMBER:  32,245
;      REFERENCE/DOCKET NUMBER:  GI 5186A
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (617) 876-1170
;      TELEFAX:  (617) 876-5851
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  428 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-050-132A-2

  Query Match             97.7%;  Score 1336.5;  DB 1;  Length 428;
  Best Local Similarity   98.4%;  Pred. No. 5.2e-128;
  Matches  253;  Conservative    2;  Mismatches    1;  Indels    1;  Gaps    1;

Qy      1 DVLEDSETWDQATGTKTFLVSQDIRDEGWETLEVSSAVKRWVRADSTTNKNKLEVTVQSH 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    173 DVLEDSETWDQATGTKTFLVSQDIRDEGWETLEVSSAVKRWVRADSTTNKNKLEVTVQSH 232

Qy     61 RESCDTLDISVPPGSKNLPFFVVFSNDRSNGTKETRLDLLKEMIGHEQETMLVKTAKNAY 120
          |||||||||||||||||||||||||||||||||||: ||||||||||||||||||||||
Db    233 RESCDTLDISVPPGSKNLPFFVVFSNDRSNGTKETRLE-LKEMIGHEQETMLVKTAKNAY 291

Qy    121 QGAGESQEEEGLDGYTAVGPLLARRKRSTGASSHCQKTSLRVNFEDIGWDSWIIAPKEYD 180
          | ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    292 QVAGESQEEEGLDGYTAVGPLLARRKRSTGASSHCQKTSLRVNFEDIGWDSWIIAPKEYD 351

Qy    181 AYECKGGCFFPLADDVTPTKHAIVQTLVHLKFPTKVGKACCVPTKLSPISILYKDDMGVP 240
          ||||||||||||||||||||||||||||||:||||||||||||||||||||||||||||
Db    352 AYECKGGCFFPLADDVTPTKHAIVQTLVHLEFPTKVGKACCVPTKLSPISILYKDDMGVP 411

Qy    241 TLKYHYEGMSVAECGCR 257
          |||||||||||||||||
Db    412 TLKYHYEGMSVAECGCR 428


RESULT 3
US-08-750-222A-2
; Sequence 2, Application US/08750222A
; Patent No. 6034061
;  GENERAL INFORMATION:
;    APPLICANT:  Rosen, Vicki A.
;    APPLICANT:  Wozney, John M.
;    APPLICANT:  Celeste, Anthony J.
;    APPLICANT:  Song, Jeffrey
;    APPLICANT:  Thies, Scott
;    TITLE OF INVENTION:  BMP-9 COMPOSITIONS
;    NUMBER OF SEQUENCES:  19
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Genetics Institute, Inc.
;      STREET:  Legal Affairs - 87 CambridgePark Drive
;      CITY:  Cambridge
;      STATE:  MA
;      COUNTRY:  US
;      ZIP:  02140
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/750,222A
;      FILING DATE:  04-DEC-1996
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/254,353
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Kapinos, Ellen J.
;      REGISTRATION NUMBER:  32,245
;      REFERENCE/DOCKET NUMBER:  GI 5186B
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (617) 876-1170
;      TELEFAX:  (617) 876-5851
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  428 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-750-222A-2

  Query Match             97.7%;  Score 1336.5;  DB 3;  Length 428;
  Best Local Similarity   98.4%;  Pred. No. 5.2e-128;
  Matches  253;  Conservative    2;  Mismatches    1;  Indels    1;  Gaps    1;

Qy      1 DVLEDSETWDQATGTKTFLVSQDIRDEGWETLEVSSAVKRWVRADSTTNKNKLEVTVQSH 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    173 DVLEDSETWDQATGTKTFLVSQDIRDEGWETLEVSSAVKRWVRADSTTNKNKLEVTVQSH 232

Qy     61 RESCDTLDISVPPGSKNLPFFVVFSNDRSNGTKETRLDLLKEMIGHEQETMLVKTAKNAY 120
          |||||||||||||||||||||||||||||||||||: ||||||||||||||||||||||
Db    233 RESCDTLDISVPPGSKNLPFFVVFSNDRSNGTKETRLE-LKEMIGHEQETMLVKTAKNAY 291

Qy    121 QGAGESQEEEGLDGYTAVGPLLARRKRSTGASSHCQKTSLRVNFEDIGWDSWIIAPKEYD 180
          | ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db    292 QVAGESQEEEGLDGYTAVGPLLARRKRSTGASSHCQKTSLRVNFEDIGWDSWIIAPKEYD 351

Qy    181 AYECKGGCFFPLADDVTPTKHAIVQTLVHLKFPTKVGKACCVPTKLSPISILYKDDMGVP 240
          ||||||||||||||||||||||||||||||:||||||||||||||||||||||||||||||
Db    352 AYECKGGCFFPLADDVTPTKHAIVQTLVHLEFPTKVGKACCVPTKLSPISILYKDDMGVP 411

Qy    241 TLKYHYEGMSVAECGCR 257
          |||||||||||||||||
Db    412 TLKYHYEGMSVAECGCR 428


RESULT 4
US-08-815-652B-2
; Sequence 2, Application US/08815652B
; Patent No. 6034062
;  GENERAL INFORMATION:
;    APPLICANT:  Wozney, John M.
;    APPLICANT:  Celeste, Anthony
;    APPLICANT:  Song, Jeffrey
;    APPLICANT:  Thies, R. Scott
;    TITLE OF INVENTION:  BMP-9 COMPOSITIONS
;    NUMBER OF SEQUENCES:  19
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Genetics Institute, Inc.
;      STREET:  Legal Affairs - 87 CambridgePark Drive
;      CITY:  Cambridge
;      STATE:  MA
;      COUNTRY:  US
;      ZIP:  02140
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/815,652B
;      FILING DATE:
;      CLASSIFICATION:  514
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Kapinos, Ellen J.
;      REGISTRATION NUMBER:  32,245
;      REFERENCE/DOCKET NUMBER:  GI 5186D
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (617) 876-1170
;      TELEFAX:  (617) 876-5851
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  428 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-815-652B-2

  Query Match             97.7%;  Score 1336.5;  DB 3;  Length 428;
  Best Local Similarity   98.4%;  Pred. No. 5.2e-128;
  Matches  253;  Conservative    2;  Mismatches    1;  Indels    1;  Gaps    1;

Qy      1 DVLEDSETWDQATGTKTFLVSQDIRDEGWETLEVSSAVKRWVRADSTTNKNKLEVTVQSH 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    173 DVLEDSETWDQATGTKTFLVSQDIRDEGWETLEVSSAVKRWVRADSTTNKNKLEVTVQSH 232

Qy     61 RESCDTLDISVPPGSKNLPFFVVFSNDRSNGTKETRLDLLKEMIGHEQETMLVKTAKNAY 120
          |||||||||||||||||||||||||||||||||: ||||||||||||||||||||
Db    233 RESCDTLDISVPPGSKNLPFFVVFSNDRSNGTKETRLE-LKEMIGHEQETMLVKTAKNAY 291

Qy    121 QGAGESQEEEGLDGYTAVGPLLARRKRSTGASSHCQKTSLRVNFEDIGWDSWIIAPKEYD 180
          | ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    292 QVAGESQEEEGLDGYTAVGPLLARRKRSTGASSHCQKTSLRVNFEDIGWDSWIIAPKEYD 351

Qy    181 AYECKGGCFFPLADDVTPTKHAIVQTLVHLKFPTKVGKACCVPTKLSPISILYKDDMGVP 240
          ||||||||||||||||||||||||||||||:||||||||||||||||||||||||||||||
Db    352 AYECKGGCFFPLADDVTPTKHAIVQTLVHLEFPTKVGKACCVPTKLSPISILYKDDMGVP 411

Qy    241 TLKYHYEGMSVAECGCR 257
          |||||||||||||||||
Db    412 TLKYHYEGMSVAECGCR 428


RESULT 5
US-08-254-353A-2
; Sequence 2, Application US/08254353A
; Patent No. 6287816
;  GENERAL INFORMATION:
;    APPLICANT:  Rosen, Vicki A.
;    APPLICANT:  Wozney, John M.
;    APPLICANT:  Celeste, Anthony J.
;    APPLICANT:  Song, Jeffrey
;    APPLICANT:  Thies, Scott
;    TITLE OF INVENTION:  BMP-9 COMPOSITIONS
;    NUMBER OF SEQUENCES:  19
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Genetics Institute, Inc.
;      STREET:  Legal Affairs - 87 CambridgePark Drive
;      CITY:  Cambridge
;      STATE:  MA
;      COUNTRY:  US
;      ZIP:  02140
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/254,353A
;      FILING DATE:
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Kapinos, Ellen J.
;      REGISTRATION NUMBER:  32,245
;      REFERENCE/DOCKET NUMBER:  GI 5186B
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (617) 876-1170
;      TELEFAX:  (617) 876-5851
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  428 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-254-353A-2

  Query Match             97.7%;  Score 1336.5;  DB 4;  Length 428;
  Best Local Similarity   98.4%;  Pred. No. 5.2e-128;
  Matches  253;  Conservative    2;  Mismatches    1;  Indels    1;  Gaps    1;

Qy      1 DVLEDSETWDQATGTKTFLVSQDIRDEGWETLEVSSAVKRWVRADSTTNKNKLEVTVQSH 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    173 DVLEDSETWDQATGTKTFLVSQDIRDEGWETLEVSSAVKRWVRADSTTNKNKLEVTVQSH 232

Qy     61 RESCDTLDISVPPGSKNLPFFVVFSNDRSNGTKETRLDLLKEMIGHEQETMLVKTAKNAY 120
          |||||||||||||||||||||||||||||||||: ||||||||||||||||||||
Db    233 RESCDTLDISVPPGSKNLPFFVVFSNDRSNGTKETRLE-LKEMIGHEQETMLVKTAKNAY 291

Qy    121 QGAGESQEEEGLDGYTAVGPLLARRKRSTGASSHCQKTSLRVNFEDIGWDSWIIAPKEYD 180
          | ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    292 QVAGESQEEEGLDGYTAVGPLLARRKRSTGASSHCQKTSLRVNFEDIGWDSWIIAPKEYD 351

Qy    181 AYECKGGCFFPLADDVTPTKHAIVQTLVHLKFPTKVGKACCVPTKLSPISILYKDDMGVP 240
          ||||||||||||||||||||||||||||||:||||||||||||||||||||||||||||||
Db    352 AYECKGGCFFPLADDVTPTKHAIVQTLVHLEFPTKVGKACCVPTKLSPISILYKDDMGVP 411

Qy    241 TLKYHYEGMSVAECGCR 257
          |||||||||||||||||
Db    412 TLKYHYEGMSVAECGCR 428
```

```
RESULT 6
PCT-US92-05374A-2
; Sequence 2, Application PC/TUS9205374A
;  GENERAL INFORMATION:
;    APPLICANT:  Wozney, John M.
;    APPLICANT:  Celeste, Anthony
;    TITLE OF INVENTION:  BMP-9 COMPOSITIONS
;    NUMBER OF SEQUENCES:  9
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Genetics Institute, Inc.
;      STREET:  Legal Affairs - 87 CambridgePark Drive
;      CITY:  Cambridge
;      STATE:  MA
;      COUNTRY:  US
;      ZIP:  02140
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/US92/05374A
;      FILING DATE:  19920625
;      CLASSIFICATION:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Kapinos, Ellen J.
;      REGISTRATION NUMBER:  32,245
;      REFERENCE/DOCKET NUMBER:  GI 5186A
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (617) 876-1170
;      TELEFAX:  (617) 876-5851
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  428 amino acids
;      TYPE:  AMINO ACID
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
PCT-US92-05374A-2

  Query Match             97.7%;  Score 1336.5;  DB 5;  Length 428;
  Best Local Similarity   98.4%;  Pred. No. 5.2e-128;
  Matches  253;  Conservative    2;  Mismatches    1;  Indels    1;  Gaps    1;

Qy       1 DVLEDSETWDQATGTKTFLVSQDIRDEGWETLEVSSAVKRWVRADSTTNKNKLEVTVQSH 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     173 DVLEDSETWDQATGTKTFLVSQDIRDEGWETLEVSSAVKRWVRADSTTNKNKLEVTVQSH 232

Qy      61 RESCDTLDISVPPGSKNLPFFVVFSNDRSNGTKETRLDLLKEMIGHEQETMLVKTAKNAY 120
           |||||||||||||||||||||||||||||||||||: ||||||||||||||||||||||
Db     233 RESCDTLDISVPPGSKNLPFFVVFSNDRSNGTKETRLE-LKEMIGHEQETMLVKTAKNAY 291

Qy     121 QGAGESQEEEGLDGYTAVGPLLARRKRSTGASSHCQKTSLRVNFEDIGWDSWIIAPKEYD 180
           | ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     292 QVAGESQEEEGLDGYTAVGPLLARRKRSTGASSHCQKTSLRVNFEDIGWDSWIIAPKEYD 351

Qy     181 AYECKGGCFFPLADDVTPTKHAIVQTLVHLKFPTKVGKACCVPTKLSPISILYKDDMGVP 240
           ||||||||||||||||||||||||||||:||||||||||||||||||||||||||||||||
Db     352 AYECKGGCFFPLADDVTPTKHAIVQTLVHLEFPTKVGKACCVPTKLSPISILYKDDMGVP 411

Qy     241 TLKYHYEGMSVAECGCR 257
           |||||||||||||||||
Db     412 TLKYHYEGMSVAECGCR 428


RESULT 7
PCT-US95-07084-2
; Sequence 2, Application PC/TUS9507084
;  GENERAL INFORMATION:
;    APPLICANT:  Rosen, Vicki A.
;    APPLICANT:  Wozney, John M.
;    APPLICANT:  Celeste, Anthony J.
;    TITLE OF INVENTION:  BMP-9 COMPOSITIONS
;    NUMBER OF SEQUENCES:  9
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Genetics Institute, Inc.
;      STREET:  Legal Affairs - 87 CambridgePark Drive
;      CITY:  Cambridge
;      STATE:  MA
;      COUNTRY:  US
;      ZIP:  02140
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/US95/07084
;      FILING DATE:
;      CLASSIFICATION:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Kapinos, Ellen J.
;      REGISTRATION NUMBER:  32,245
;      REFERENCE/DOCKET NUMBER:  GI 5186C-PCT
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (617) 876-1210
;      TELEFAX:  (617) 876-5851
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  428 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
PCT-US95-07084-2

  Query Match             97.7%;  Score 1336.5;  DB 5;  Length 428;
  Best Local Similarity   98.4%;  Pred. No. 5.2e-128;
  Matches  253;  Conservative    2;  Mismatches    1;  Indels    1;  Gaps    1;

Qy       1 DVLEDSETWDQATGTKTFLVSQDIRDEGWETLEVSSAVKRWVRADSTTNKNKLEVTVQSH 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     173 DVLEDSETWDQATGTKTFLVSQDIRDEGWETLEVSSAVKRWVRADSTTNKNKLEVTVQSH 232

Qy      61 RESCDTLDISVPPGSKNLPFFVVFSNDRSNGTKETRLDLLKEMIGHEQETMLVKTAKNAY 120
           |||||||||||||||||||||||||||||||||||: ||||||||||||||||||||||
Db     233 RESCDTLDISVPPGSKNLPFFVVFSNDRSNGTKETRLE-LKEMIGHEQETMLVKTAKNAY 291

Qy     121 QGAGESQEEEGLDGYTAVGPLLARRKRSTGASSHCQKTSLRVNFEDIGWDSWIIAPKEYD 180
           | ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     292 QVAGESQEEEGLDGYTAVGPLLARRKRSTGASSHCQKTSLRVNFEDIGWDSWIIAPKEYD 351

Qy     181 AYECKGGCFFPLADDVTPTKHAIVQTLVHLKFPTKVGKACCVPTKLSPISILYKDDMGVP 240
           ||||||||||||||||||||||||||||:||||||||||||||||||||||||||||||||
Db     352 AYECKGGCFFPLADDVTPTKHAIVQTLVHLEFPTKVGKACCVPTKLSPISILYKDDMGVP 411

Qy     241 TLKYHYEGMSVAECGCR 257
           |||||||||||||||||
Db     412 TLKYHYEGMSVAECGCR 428


RESULT 8
US-08-065-844A-2
; Sequence 2, Application US/08065844A
; Patent No. 6333168
;  GENERAL INFORMATION:
;    APPLICANT:  Jessell, Thomas M.
;    APPLICANT:  Basler, Konrad
;    APPLICANT:  Yomada, Toshiya
;    TITLE OF INVENTION:  CLONING, EXPRESSION AND USES OF
;    TITLE OF INVENTION:  DORSALIN-1
;    NUMBER OF SEQUENCES:  18
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Cooper & Dunham
;      STREET:  30 Rockefeller Plaza
;      CITY:  New York
;      STATE:  New York
;      COUNTRY:  United States of America
```

```
;      ZIP:  10112
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/065,844A
;      FILING DATE:  19930520
;      CLASSIFICATION:  424
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  White, John P.
;      REGISTRATION NUMBER:  28,678
;      REFERENCE/DOCKET NUMBER:  0576/40314
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (212) 977-9550
;      TELEFAX:  (212) 664-0525
;      TELEX:  422523 COOP UI
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  427 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-065-844A-2

  Query Match             59.1%;  Score 808.5;  DB 4;  Length 427;
  Best Local Similarity   61.9%;  Pred. No. 3.4e-74;
  Matches  161;  Conservative   31;  Mismatches   59;  Indels    9;  Gaps    5;

Qy        1 DVLEDSETWDQATGTKTFLVSQDIRDEGWETLEVSSAVKRWVRADSTTNKNKLEVTVQSH 60
            ||| | : |:    ||: ||| |:| ||| ||||||||||:||    |||||| ::|
Db      174 DVL-DGDHWENKESTKSLLVSHSIQDCGWEMFEVSSAVKRWVKADKMKTKNKLEVVIESK 232

Qy       61 RES---CDTLDISVPPGSKNLPFFVVFSNDRSNGTKETRLDLLKEMIGHEQETMLVKTAK 117
              |   |  |||:|   :|||| :|||||||||||||::: |:||| |||::| |  |
Db      233 DLSGFPCGKLDITVTHDTKNLPLLIVFSNDRSNGTKETKVE-LREMIVHEQESVLNKLGK 291

Qy      118 NAYQGAGESQEEEGLDGYTAVGPLLARRKRSTGASSHCQKTSLRVNFEDIGWDSWIIAPK 177
            |      | :||: :          :| ||| || :||::||| |||::|||||||||||
Db      292 NDSSSEEEQREEKAI---ARPRQHSSRSKRSIGA-NHCRRTSLHVNFKEIGWDSWIIAPK 347

Qy      178 EYDAYECKGGCFFPLADDVTPTKHAIVQTLVHLKFPTKVGKACCVPTKLSPISILYKDDM 237
            :|:|:|||||||||| |:|||||||||||||||: | |  |||||||||  ||||||||
Db      348 DYEAFECKGGCFFPLTDNVTPTKHAIVQTLVHLQNPKKASKACCVPTKLDAISILYKDDA 407

Qy      238 GVPTLKYHYEGMSVAECGCR 257
            ||||| |:|||| |||||||
Db      408 GVPTLIYNYEGMKVAECGCR 427


RESULT 9
US-08-050-132A-9
; Sequence 9, Application US/08050132A
; Patent No. 5661007
;  GENERAL INFORMATION:
;    APPLICANT:  Wozney, John M.
;    APPLICANT:  Celeste, Anthony
;    TITLE OF INVENTION:  BMP-9 COMPOSITIONS
;    NUMBER OF SEQUENCES:  19
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Genetics Institute, Inc.
;      STREET:  Legal Affairs - 87 CambridgePark Drive
;      CITY:  Cambridge
;      STATE:  MA
;      COUNTRY:  US
;      ZIP:  02140
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/050,132A
;      FILING DATE:
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Kapinos, Ellen J.
;      REGISTRATION NUMBER:  32,245
;      REFERENCE/DOCKET NUMBER:  GI 5186A
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (617) 876-1170
;      TELEFAX:  (617) 876-5851
;  INFORMATION FOR SEQ ID NO:  9:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  150 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-050-132A-9

  Query Match             46.3%;  Score 634;  DB 1;  Length 150;
  Best Local Similarity   92.7%;  Pred. No. 4.5e-57;
  Matches  114;  Conservative    3;  Mismatches    6;  Indels    0;  Gaps    0;

Qy      135 YTAVGPLLARRKRSTGASSHCQKTSLRVNFEDIGWDSWIIAPKEYDAYECKGGCFFPLAD 194
            : | |  ||||||| || ||||||||||||||||||||||||||||:|||||||||||||
Db       28 HVAAGSTLARRKRSAGAGSHCQKTSLRVNFEDIGWDSWIIAPKEYEAYECKGGCFFPLAD 87

Qy      195 DVTPTKHAIVQTLVHLKFPTKVGKACCVPTKLSPISILYKDDMGVPTLKYHYEGMSVAEC 254
            ||||||||||||||||||||||||||||||||||||:|||||||||||||||||||||||
Db       88 DVTPTKHAIVQTLVHLKFPTKVGKACCVPTKLSPISVLYKDDMGVPTLKYHYEGMSVAEC 147

Qy      255 GCR 257
            |||
Db      148 GCR 150


RESULT 10
US-08-750-222A-9
; Sequence 9, Application US/08750222A
; Patent No. 6034061
;  GENERAL INFORMATION:
;    APPLICANT:  Rosen, Vicki A.
;    APPLICANT:  Wozney, John M.
;    APPLICANT:  Celeste, Anthony J.
;    APPLICANT:  Song, Jeffrey
;    APPLICANT:  Thies, Scott
;    TITLE OF INVENTION:  BMP-9 COMPOSITIONS
;    NUMBER OF SEQUENCES:  19
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Genetics Institute, Inc.
;      STREET:  Legal Affairs - 87 CambridgePark Drive
;      CITY:  Cambridge
;      STATE:  MA
;      COUNTRY:  US
;      ZIP:  02140
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/750,222A
;      FILING DATE:  04-DEC-1996
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/254,353
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Kapinos, Ellen J.
;      REGISTRATION NUMBER:  32,245
;      REFERENCE/DOCKET NUMBER:  GI 5186B
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (617) 876-1170
;      TELEFAX:  (617) 876-5851
```

```
;  INFORMATION FOR SEQ ID NO:  9:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  150 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-750-222A-9

  Query Match             46.3%;  Score 634;  DB 3;  Length 150;
  Best Local Similarity   92.7%;  Pred. No. 4.5e-57;
  Matches  114;  Conservative    3;  Mismatches    6;  Indels    0;  Gaps    0;

Qy    135 YTAVGPLLARRKRSTGASSHCQKTSLRVNFEDIGWDSWIIAPKEYDAYECKGGCFFPLAD 194
          : | |  ||||||| || ||||||||||||||||||||||||||:|||||||||||||||
Db     28 HVAAGSTLARRKRSAGAGSHCQKTSLRVNFEDIGWDSWIIAPKEYEAYECKGGCFFPLAD 87

Qy    195 DVTPTKHAIVQTLVHLKFPTKVGKACCVPTKLSPISILYKDDMGVPTLKYHYEGMSVAEC 254
          ||||||||||||||||||||||||||||||||||||:||||||||||||||||||||||
Db     88 DVTPTKHAIVQTLVHLKFPTKVGKACCVPTKLSPISVLYKDDMGVPTLKYHYEGMSVAEC 147

Qy    255 GCR 257
          |||
Db    148 GCR 150


RESULT 11
US-08-815-652B-9
; Sequence 9, Application US/08815652B
; Patent No. 6034062
;  GENERAL INFORMATION:
;    APPLICANT:  Wozney, John M.
;    APPLICANT:  Celeste, Anthony
;    APPLICANT:  Song, Jeffrey
;    APPLICANT:  Thies, R. Scott
;    TITLE OF INVENTION:  BMP-9 COMPOSITIONS
;    NUMBER OF SEQUENCES:  19
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Genetics Institute, Inc.
;      STREET:  Legal Affairs - 87 CambridgePark Drive
;      CITY:  Cambridge
;      STATE:  MA
;      COUNTRY:  US
;      ZIP:  02140
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/815,652B
;      FILING DATE:
;      CLASSIFICATION:  514
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Kapinos, Ellen J.
;      REGISTRATION NUMBER:  32,245
;      REFERENCE/DOCKET NUMBER:  GI 5186D
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (617) 876-1170
;      TELEFAX:  (617) 876-5851
;  INFORMATION FOR SEQ ID NO:  9:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  150 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-815-652B-9

  Query Match             46.3%;  Score 634;  DB 3;  Length 150;
  Best Local Similarity   92.7%;  Pred. No. 4.5e-57;
  Matches  114;  Conservative    3;  Mismatches    6;  Indels    0;  Gaps    0;

Qy    135 YTAVGPLLARRKRSTGASSHCQKTSLRVNFEDIGWDSWIIAPKEYDAYECKGGCFFPLAD 194
          : | |  ||||||| || ||||||||||||||||||||||||||:|||||||||||||||
Db     28 HVAAGSTLARRKRSAGAGSHCQKTSLRVNFEDIGWDSWIIAPKEYEAYECKGGCFFPLAD 87

Qy    195 DVTPTKHAIVQTLVHLKFPTKVGKACCVPTKLSPISILYKDDMGVPTLKYHYEGMSVAEC 254
          ||||||||||||||||||||||||||||||||||||:||||||||||||||||||||||
Db     88 DVTPTKHAIVQTLVHLKFPTKVGKACCVPTKLSPISVLYKDDMGVPTLKYHYEGMSVAEC 147

Qy    255 GCR 257
          |||
Db    148 GCR 150


RESULT 12
US-08-254-353A-9
; Sequence 9, Application US/08254353A
; Patent No. 6287816
;  GENERAL INFORMATION:
;    APPLICANT:  Rosen, Vicki A.
;    APPLICANT:  Wozney, John M.
;    APPLICANT:  Celeste, Anthony J.
;    APPLICANT:  Song, Jeffrey
;    APPLICANT:  Thies, Scott
;    TITLE OF INVENTION:  BMP-9 COMPOSITIONS
;    NUMBER OF SEQUENCES:  19
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Genetics Institute, Inc.
;      STREET:  Legal Affairs - 87 CambridgePark Drive
;      CITY:  Cambridge
;      STATE:  MA
;      COUNTRY:  US
;      ZIP:  02140
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/254,353A
;      FILING DATE:
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Kapinos, Ellen J.
;      REGISTRATION NUMBER:  32,245
;      REFERENCE/DOCKET NUMBER:  GI 5186B
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (617) 876-1170
;      TELEFAX:  (617) 876-5851
;  INFORMATION FOR SEQ ID NO:  9:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  150 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-254-353A-9

  Query Match             46.3%;  Score 634;  DB 4;  Length 150;
  Best Local Similarity   92.7%;  Pred. No. 4.5e-57;
  Matches  114;  Conservative    3;  Mismatches    6;  Indels    0;  Gaps    0;

Qy    135 YTAVGPLLARRKRSTGASSHCQKTSLRVNFEDIGWDSWIIAPKEYDAYECKGGCFFPLAD 194
          : | |  ||||||| || ||||||||||||||||||||||||||:|||||||||||||||
Db     28 HVAAGSTLARRKRSAGAGSHCQKTSLRVNFEDIGWDSWIIAPKEYEAYECKGGCFFPLAD 87

Qy    195 DVTPTKHAIVQTLVHLKFPTKVGKACCVPTKLSPISILYKDDMGVPTLKYHYEGMSVAEC 254
          ||||||||||||||||||||||||||||||||||||:||||||||||||||||||||||
Db     88 DVTPTKHAIVQTLVHLKFPTKVGKACCVPTKLSPISVLYKDDMGVPTLKYHYEGMSVAEC 147

Qy    255 GCR 257
          |||
Db    148 GCR 150


RESULT 13
PCT-US92-05374A-9
```

```
; Sequence 9, Application PC/TUS9205374A
;  GENERAL INFORMATION:
;     APPLICANT:  Wozney, John M.
;     APPLICANT:  Celeste, Anthony
;     TITLE OF INVENTION:  BMP-9 COMPOSITIONS
;     NUMBER OF SEQUENCES:  9
;     CORRESPONDENCE ADDRESS:
;       ADDRESSEE:  Genetics Institute, Inc.
;       STREET:  Legal Affairs - 87 CambridgePark Drive
;       CITY:  Cambridge
;       STATE:  MA
;       COUNTRY:  US
;       ZIP:  02140
;     COMPUTER READABLE FORM:
;       MEDIUM TYPE:  Floppy disk
;       COMPUTER:  IBM PC compatible
;       OPERATING SYSTEM:  PC-DOS/MS-DOS
;       SOFTWARE:  PatentIn Release #1.0, Version #1.25
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  PCT/US92/05374A
;       FILING DATE:  19920625
;       CLASSIFICATION:
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  Kapinos, Ellen J.
;       REGISTRATION NUMBER:  32,245
;       REFERENCE/DOCKET NUMBER:  GI 5186A
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  (617) 876-1170
;       TELEFAX:  (617) 876-5851
;   INFORMATION FOR SEQ ID NO:  9:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  150 amino acids
;       TYPE:  AMINO ACID
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  protein
PCT-US92-05374A-9

  Query Match             46.3%;  Score 634;  DB 5;  Length 150;
  Best Local Similarity   92.7%;  Pred. No. 4.5e-57;
  Matches  114;  Conservative    3;  Mismatches    6;  Indels    0;  Gaps    0;

Qy     135 YTAVGPLLARRKRSTGASSHCQKTSLRVNFEDIGWDSWIIAPKEYDAYECKGGCFFPLAD 194
           : | |  ||||||| || ||||||||||||||||||||||||||||:|||||||||||||||
Db      28 HVAAGSTLARRKRSAGAGSHCQKTSLRVNFEDIGWDSWIIAPKEYEAYECKGGCFFPLAD 87

Qy     195 DVTPTKHAIVQTLVHLKFPTKVGKACCVPTKLSPISILYKDDMGVPTLKYHYEGMSVAEC 254
           ||||||||||||||||||||||||||||||||||:||||||||||||||||||||||||||
Db      88 DVTPTKHAIVQTLVHLKFPTKVGKACCVPTKLSPISVLYKDDMGVPTLKYHYEGMSVAEC 147

Qy     255 GCR 257
           |||
Db     148 GCR 150


RESULT 14
PCT-US95-07084-9
; Sequence 9, Application PC/TUS9507084
;  GENERAL INFORMATION:
;     APPLICANT:  Rosen, Vicki A.
;     APPLICANT:  Wozney, John M.
;     APPLICANT:  Celeste, Anthony J.
;     TITLE OF INVENTION:  BMP-9 COMPOSITIONS
;     NUMBER OF SEQUENCES:  9
;     CORRESPONDENCE ADDRESS:
;       ADDRESSEE:  Genetics Institute, Inc.
;       STREET:  Legal Affairs - 87 CambridgePark Drive
;       CITY:  Cambridge
;       STATE:  MA
;       COUNTRY:  US
;       ZIP:  02140
;     COMPUTER READABLE FORM:
;       MEDIUM TYPE:  Floppy disk
;       COMPUTER:  IBM PC compatible
;       OPERATING SYSTEM:  PC-DOS/MS-DOS
;       SOFTWARE:  PatentIn Release #1.0, Version #1.25
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  PCT/US95/07084
;       FILING DATE:
;       CLASSIFICATION:
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  Kapinos, Ellen J.
;       REGISTRATION NUMBER:  32,245
;       REFERENCE/DOCKET NUMBER:  GI 5186C-PCT
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  (617) 876-1210
;       TELEFAX:  (617) 876-5851
;   INFORMATION FOR SEQ ID NO:  9:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  150 amino acids
;       TYPE:  amino acid
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  protein
PCT-US95-07084-9

  Query Match             46.3%;  Score 634;  DB 5;  Length 150;
  Best Local Similarity   92.7%;  Pred. No. 4.5e-57;
  Matches  114;  Conservative    3;  Mismatches    6;  Indels    0;  Gaps    0;

Qy     135 YTAVGPLLARRKRSTGASSHCQKTSLRVNFEDIGWDSWIIAPKEYDAYECKGGCFFPLAD 194
           : | |  ||||||| || ||||||||||||||||||||||||||||:|||||||||||||||
Db      28 HVAAGSTLARRKRSAGAGSHCQKTSLRVNFEDIGWDSWIIAPKEYEAYECKGGCFFPLAD 87

Qy     195 DVTPTKHAIVQTLVHLKFPTKVGKACCVPTKLSPISILYKDDMGVPTLKYHYEGMSVAEC 254
           ||||||||||||||||||||||||||||||||||:||||||||||||||||||||||||||
Db      88 DVTPTKHAIVQTLVHLKFPTKVGKACCVPTKLSPISVLYKDDMGVPTLKYHYEGMSVAEC 147

Qy     255 GCR 257
           |||
Db     148 GCR 150


RESULT 15
US-08-247-908A-2
; Sequence 2, Application US/08247908A
; Patent No. 5637480
;  GENERAL INFORMATION:
;     APPLICANT:  CELESTE, Anthony J
;     APPLICANT:  WOZNEY, John
;     TITLE OF INVENTION:  BMP-10 COMPOSITIONS
;     NUMBER OF SEQUENCES:  12
;     CORRESPONDENCE ADDRESS:
;       ADDRESSEE:  GENETICS INSTITUTE, INC.
;       STREET:  87 CambridgePark Drive
;       CITY:  Cambridge
;       STATE:  MA
;       COUNTRY:  USA
;       ZIP:  02144
;     COMPUTER READABLE FORM:
;       MEDIUM TYPE:  Floppy disk
;       COMPUTER:  IBM PC compatible
;       OPERATING SYSTEM:  PC-DOS/MS-DOS
;       SOFTWARE:  PatentIn Release #1.0, Version #1.25
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/08/247,908A
;       FILING DATE:
;       CLASSIFICATION:  435
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  LAZAR, Steven R.
;       REGISTRATION NUMBER:  32,618
;       REFERENCE/DOCKET NUMBER:  GI 5206-CIP
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  617 876-1170 x8260
;       TELEFAX:  617 876-5851
;   INFORMATION FOR SEQ ID NO:  2:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  312 amino acids
```

```
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-247-908A-2

  Query Match            38.0%;  Score 519.5;  DB 1;  Length 312;
  Best Local Similarity  42.8%;  Pred. No. 6.1e-45;
  Matches  119;  Conservative   43;  Mismatches   73;  Indels   43;  Gaps   10;

Qy       1 DVLEDSETWDQATGTKTFLVSQDI--RDEGWETLEVSSAVKRWVRADSTTNKNKLEVTVQ 58
           :|||  |   :       ||| :|   :  ||| :|: |:: | :: |:|  ::||| ::
Db      57 EVLESKED-HEGERNMLVLVSGEIYGTNSEWETFDVTDAIRHWQKSGSST--HQLEVHIE 113

Qy      59 SHRESCDT-----LDISVPPGSKNLPFFVVFSNDRSNGTKETRLDLLKEMIGHEQETMLV 113
           |  |  ||     |:|     :|: |  ||||:|:|  ::: | : | ||| |||
Db     114 SKHEMEDTLGRGQLEIDTSARNKHDPLLVVFSDDQS--SEKERKEELDEMIAHEQ----- 166

Qy     114 KTAKNAYQGAGESQEEEGLDGYTAVGP----LLARRK----------RSTGASSHCQKTS 159
                       :   ||||:  ||    ||  |           |     ::|::|
Db     167 ----------FPEMDNLDLDGYSN-GPGEEALLQMRSNIIYDSTARIRRNAKGNYCKRTP 215

Qy     160 LRVNFEDIGWDSWIIAPKEYDAYECKGGCFFPLADDVTPTKHAIVQTLVHLKFPTKVGKA 219
           | ::|::||||||||||  |:||||:| | :|||: :|||||||:| |||||   |  ||
Db     216 LYIDFKEIGWDSWIIAPPGYEAYECRGVCNYPLAEHLTPTKHAIIQALVHLKNSQKASKA 275

Qy     220 CCVPTKLSPISILYKDDMGVPTLKYHYEGMSVAECGCR 257
           ||||||| ||||||  | || | |: ||||:|:|||||
Db     276 CCVPTKLEPISILYL-DKGVVTYKFKYEGMAVSECGCR 312


Search completed: April  2, 2003, 14:53:07
Job time : 30 secs
```

GenCore version 5.1.4_p5_4578

OM protein - protein search, using sw model

Run on: April 2, 2003, 14:46:20 ; Search time 86 Seconds
(without alignments)
615.745 Million cell updates/sec

Title: US-10-002-278-9
Perfect score: 1368
Sequence: 1 DVLEDSETWDQATGTKTFLV.....GVPTLKYHYEGMSVAECGCR 257

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SPTREMBL_21:*
1: sp_archea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_rvirus:*
16: sp_bacteriap:*
17: sp_archeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	357.5	26.1	411	5	Q9U418	Q9u418 branchiosto
2	352	25.7	361	5	O96504	O96504 branchiosto
3	346	25.3	405	5	Q9U5E8	Q9u5e8 ptychodera
4	341	24.9	337	6	Q9MZV5	Q9mzv5 canis famil
5	335.5	24.5	399	13	Q9W753	Q9w753 xenopus lae
6	333	24.3	614	5	P91720	P91720 drosophila
7	329.5	24.1	261	13	Q9W6C0	Q9w6c0 brachydanio
8	326.5	23.9	391	13	Q8UVQ8	Q8uvq8 brachydanio
9	326.5	23.9	400	13	O73818	O73818 xenopus lae
10	326.5	23.9	400	13	Q91703	Q91703 xenopus lae
11	326.5	23.9	509	5	Q8WS99	Q8ws99 archaster t
12	325.5	23.8	391	13	Q8UVQ2	Q8uvq2 brachydanio
13	325	23.8	588	5	Q9VQC6	Q9vqc6 drosophila
14	320.5	23.4	438	5	Q9GT26	Q9gt26 anopheles s
15	320	23.4	391	13	Q90YJ3	Q90yj3 brachydanio
16	319.5	23.4	400	13	Q90YD6	Q90yd6 xenopus tro
17	319.5	23.4	417	5	Q9XYQ7	Q9xyq7 lytechinus
18	319	23.3	204	5	Q9XZ69	Q9xz69 tripneustes
19	319	23.3	257	13	O42303	O42303 brachydanio
20	318	23.2	364	13	Q9PVK1	Q9pvk1 gallus gall
21	318	23.2	400	13	O57574	O57574 brachydanio
22	316	23.1	411	13	O57573	O57573 brachydanio
23	316	23.1	411	13	O93369	O93369 brachydanio
24	315.5	23.1	412	13	O12938	O12938 brachydanio
25	315	23.0	289	5	Q9XYQ8	Q9xyq8 strongyloce
26	314	23.0	398	13	Q90YD7	Q90yd7 xenopus tro
27	313	22.9	178	5	Q25211	Q25211 junonia coe
28	311	22.7	400	13	O13107	O13107 brachydanio
29	311	22.7	411	13	O13108	O13108 brachydanio
30	310.5	22.7	277	13	Q90Y82	Q90y82 lampetra ja
31	310.5	22.7	301	5	O97390	O97390 crassostrea
32	310.5	22.7	313	13	Q91403	Q91403 gallus gall
33	306	22.4	432	13	Q9PTF9	Q9ptf9 brachydanio
34	306	22.4	453	13	P87373	P87373 gallus gall
35	305	22.3	126	13	O93573	O93573 gallus gall
36	304.5	22.3	500	13	Q9W6G0	Q9w6g0 gallus gall
37	302.5	22.1	398	13	Q9I8T6	Q9i8t6 gallus gall
38	301.5	22.0	443	5	O76851	O76851 halocynthia
39	300.5	22.0	390	13	Q91597	Q91597 xenopus lae
40	298.5	21.8	182	13	Q90Y81	Q90y81 lampetra ja
41	298	21.8	424	13	Q9YGH7	Q9ygh7 xenopus lae
42	296	21.6	386	13	O13109	O13109 brachydanio
43	294.5	21.5	294	6	Q9BDW9	Q9bdw9 macaca fasc
44	294.5	21.5	447	6	Q9BDW8	Q9bdw8 cercopithec
45	293.5	21.5	413	13	Q9DGN4	Q9dgn4 xenopus lae

ALIGNMENTS

RESULT 1
Q9U418
ID Q9U418 PRELIMINARY; PRT; 411 AA.
AC Q9U418;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Bone morphogenetic protein 2/4.
GN BMP2/4.
OS Branchiostoma belcheri (Amphoxius).
OC Eukaryota; Metazoa; Chordata; Cephalochordata; Branchiostomidae;
OC Branchiostoma.
OX NCBI_TaxID=7741;
RN [1]
RP SEQUENCE FROM N.A.
RA Yasui K., Saiga H., Uemura M., Semba I.;
RT "Early body formation and expression pattern of genes encoding
RT secreted proteins, BbBMP2/4, BbWnt7, and BbWnt8 in lancelets.";
RL Submitted (NOV-1999) to the EMBL/GenBank/DDBJ databases.
CC -!- SIMILARITY: BELONGS TO THE TGF-BETA FAMILY.
DR EMBL; AF206325; AAF19841.1; -.
DR HSSP; P12643; 3BMP.
DR InterPro; IPR002405; Inhibin_alpha.
DR InterPro; IPR001839; TGFb.
DR InterPro; IPR001111; TGFb_N.
DR Pfam; PF00019; TGF-beta; 1.
DR Pfam; PF00688; TGFb_propeptide; 1.
DR PRINTS; PR00669; INHIBINA.
DR ProDom; PD000357; TGFb; 1.
DR SMART; SM00204; TGFB; 1.
DR PROSITE; PS00250; TGF_BETA_1; 1.
KW Glycoprotein.
SQ SEQUENCE 411 AA; 46929 MW; F579898060F18355 CRC64;

Query Match 26.1%; Score 357.5; DB 5; Length 411;
Best Local Similarity 35.5%; Pred. No. 2.2e-25;
Matches 91; Conservative 29; Mismatches 87; Indels 49; Gaps 8;

Qy 13 TGTKTFLVSQ---DIRDEGWETLEVSSAVKRWVRADSTTNKNKLEVTVQSHRESCDTLDI 69

```
          | | | |:     |:|:  ||: :| ||| :|   :|      ||| | | :
Db    194 TDTITRLLDTKLVDVRNSSWESFDVRSAVTKW--KNSPERNYGLEVEVVSPKRG------ 245

Qy     70 SVPPGSKNLPFFVVFSNDRSNGTKETRLDLLKEMIGHEQETMLVKTAKNAYQGAGESQEE 129
                      ||      : | :|       | :  :|  |      | | |
Db    246 ------------ALSNHHVRLRRSTDMD--DHAWQHRRPLLLTYTD----DGKGSSNSN 286

Qy    130 EGLDGYTAVGPLLARRKRSTGASSH--------CQKTSLRVNFEDIGWDSWIIAPKEYDA 181
                    : :|:||: |   |         |:: || |:| |:||: ||:||  | |
Db    287 R----------VASRQKRANGRKKHQRRRLKANCRRHSLYVDFSDVGWNDWIVAPPGYQA 336

Qy    182 YECKGGCFFPLADDVTPTKHAIVQTLVHLKFPTKVGKACCVPTKLSPISILYKDDMGVPT 241
          | | | |||||  :  | ||||||||:   |  | ||||||| |||||:|| ::
Db    337 YYCHGECPFPLADHLNSTNHAIVQTLVNSVNPLAVPKACCVPTDLSPISMLYLNENDQVV 396

Qy    242 LKYHYEGMSVAECGCR 257
          || :|: | |  ||||
Db    397 LK-NYQDMVVEGCGCR 411


RESULT 2
O96504
ID   O96504      PRELIMINARY;      PRT;   361 AA.
AC   O96504;
DT   01-MAY-1999 (TrEMBLrel. 10, Created)
DT   01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT   01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE   Bone morphogenetic protein 2/4.
GN   AMPHIBMP2/4.
OS   Branchiostoma floridae (Florida lancelet) (Amphioxus).
OC   Eukaryota; Metazoa; Chordata; Cephalochordata; Branchiostomidae;
OC   Branchiostoma.
OX   NCBI_TaxID=7739;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=98401944; PubMed=9733108;
RA   Panopoulou G.D., Clark M.D., Holland L.Z., Lehrach H., Holland N.D.;
RT   "AmphiBMP2/4, an amphioxus bone morphogenetic protein closely related
RT   to Drosophila decapentaplegic and vertebrate BMP2 and BMP4: insights
RT   into evolution of dorsoventral axis specification.";
RL   Dev. Dyn. 213:130-139(1998).
CC   -!- SIMILARITY: BELONGS TO THE TGF-BETA FAMILY.
DR   EMBL; AF068750; AAC97488.1; -.
DR   HSSP; P12643; 3BMP.
DR   InterPro; IPR002405; Inhibin_alpha.
DR   InterPro; IPR001839; TGFb.
DR   InterPro; IPR001111; TGFb_N.
DR   Pfam; PF00019; TGF-beta; 1.
DR   Pfam; PF00688; TGFb_propeptide; 1.
DR   PRINTS; PR00669; INHIBINA.
DR   ProDom; PD000357; TGFb; 1.
DR   SMART; SM00204; TGFB; 1.
DR   PROSITE; PS00250; TGF_BETA_1; 1.
KW   Glycoprotein.
SQ   SEQUENCE   361 AA;  41517 MW;  09FF5FE6C9785DD1 CRC64;

  Query Match           25.7%;  Score 352;  DB 5;  Length 361;
  Best Local Similarity   35.3%;  Pred. No. 6e-25;
  Matches   90;  Conservative   31;  Mismatches   86;  Indels   48;  Gaps    8;

Qy     13 TGTKTFLVSQ---DIRDEGWETLEVSSAVKRWVRADSTTNKNKLEVTVQSHRESCDTLDI 69
          | | | |:     |:|:  ||: :| ||| :|   :|      ||| | | :
Db    145 TDTITRLLDTKLVDVRNSSWESFDVRSAVTKW--KNSPERNYGLEVEVVSPKRG------ 196

Qy     70 SVPPGSKNLPFFVVFSNDRSNGTKETRLDLLKEMIGHEQETMLVKTAKNAYQGAGESQEE 129
                      ||      : | :|       | :  :|  |      | | |
Db    197 ------------ALSNHHVRLRRSTDMD--DHSWQHRRPLLLTYTD----DGKGSSNSN 237

Qy    130 EGLDGYTAVGPLLARRKRSTG-------ASSHCQKTSLRVNFEDIGWDSWIIAPKEYDAY 182
                    : :|:||: |         ::|:: || |:| |:||: ||:||  | ||
Db    238 R----------VASRQKRANGRKKQRRRLKANCRRHSLYVDFSDVGWNDWIVAPPGYQAY 287

Qy    183 ECKGGCFFPLADDVTPTKHAIVQTLVHLKFPTKVGKACCVPTKLSPISILYKDDMGVPTL 242
          | | | |||||  :  | ||||||||:   |  | ||||||| |||||:|| ::     |
Db    288 YCHGECPFPLADHLNSTNHAIVQTLVNSVNPLAVPKACCVPTDLSPISMLYLNENDQVVL 347

Qy    243 KYHYEGMSVAECGCR 257
          | :|: | |  ||||
Db    348 K-NYQDMVVEGCGCR 361


RESULT 3
Q9U5E8
ID   Q9U5E8      PRELIMINARY;      PRT;   405 AA.
AC   Q9U5E8;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE   Pf-BMP2/4.
GN   PF-BMP2/4.
OS   Ptychodera flava.
OC   Eukaryota; Metazoa; Hemichordata; Enteropneusta; Ptychoderidae;
OC   Ptychodera.
OX   NCBI_TaxID=63121;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Harada Y.;
RL   Submitted (JUN-1999) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE FROM N.A.
RA   Okai N., Taguchi S., Tagawa K., Satoh N.;
RT   "Developmental expression of hemichordate orthologs of BMP-4, otx and
RT   dlx: insights into deuterostome archetype and evolution of chordate
RT   body plan.";
RL   Submitted (JAN-2000) to the EMBL/GenBank/DDBJ databases.
CC   -!- SIMILARITY: BELONGS TO THE TGF-BETA FAMILY.
DR   EMBL; AB028219; BAA89012.1; -.
DR   HSSP; P12643; 3BMP.
DR   InterPro; IPR002405; Inhibin_alpha.
DR   InterPro; IPR001839; TGFb.
DR   InterPro; IPR001111; TGFb_N.
DR   Pfam; PF00019; TGF-beta; 1.
DR   Pfam; PF00688; TGFb_propeptide; 1.
DR   PRINTS; PR00669; INHIBINA.
DR   ProDom; PD000357; TGFb; 1.
DR   SMART; SM00204; TGFB; 1.
DR   PROSITE; PS00250; TGF_BETA_1; 1.
KW   Glycoprotein.
SQ   SEQUENCE   405 AA;  45936 MW;  D2EB2D2C7560ED13 CRC64;

  Query Match           25.3%;  Score 346;  DB 5;  Length 405;
  Best Local Similarity   31.1%;  Pred. No. 2.6e-24;
  Matches   76;  Conservative   43;  Mismatches   71;  Indels   54;  Gaps    6;

Qy     23 DIRDEGWETLEVSSAVKRWVRADSTTNKNKLEVT------VQSHRESC---DTLDISVPP 73
          |||:  ||: ::  || :|  :    :  ::|:|      :  |::      : |::
Db    207 DIRNSSWESFDIRPAVAKWKASQEENHGVEVELTEVQNSQISPHKDHVRLRRSSDLAASE 266

Qy     74 GSKNLPFFVVFSNDRSNGTKETRLDLLKEMIGHEQETMLVKTAKNAYQGAGESQEEEGLD 133
            :  |  : :::|   | : ||                    : :|:
Db    267 WQRQRPLLITYTDD---GKRPTR-----------------SKRNS-------------- 291

Qy    134 GYTAVGPLLARRKRSTGASSHCQKTSLRVNFEDIGWDSWIIAPKEYDAYECKGGCFFPLA 193
                    :|       :|:: || |:| |:||: ||:||  |:|: | | | ||||
Db    292 ---------ERKKGGRKLKPNCRRRSLYVDFSDVGWNDWIVAPPGYNAFYCDGECPFPLA 342

Qy    194 DDVTPTKHAIVQTLVHLKFPTKVGKACCVPTKLSPISILYKDDMGVPTLKYHYEGMSVAE 253
          | :  | |||||||||     : | :||||||:|||||:|| |:     || :|: | |
Db    343 DHLNSTNHAIVQTLVHSVKASAVPQACCVPTELSPISMLYLDEYDKVILK-NYQEMVVEG 401

Qy    254 CGCR 257
          ||||
Db    402 CGCR 405


RESULT 4
```

```
Q9MZV5
ID   Q9MZV5      PRELIMINARY;      PRT;   337 AA.
AC   Q9MZV5;
DT   01-OCT-2000 (TrEMBLrel. 15, Created)
DT   01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT   01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE   Bone morphogenetic protein 4 (Fragment).
GN   BMP4.
OS   Canis familiaris (Dog).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
OX   NCBI_TaxID=9615;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Johanson J.A., Breen M., Lepine A., Murphy K.E.;
RT   "Identification and chromosomal localization of the gene encoding
RT   canine bone morphogenetic protein 4 (bmp4).";
RL   Submitted (MAR-1999) to the EMBL/GenBank/DDBJ databases.
CC   -!- SIMILARITY: BELONGS TO THE TGF-BETA FAMILY.
DR   EMBL; AF136233; AAF82188.1; -.
DR   HSSP; P12643; 3BMP.
DR   InterPro; IPR001839; TGFb.
DR   InterPro; IPR001111; TGFb_N.
DR   Pfam; PF00019; TGF-beta; 1.
DR   Pfam; PF00688; TGFb_propeptide; 1.
DR   ProDom; PD000357; TGFb; 1.
DR   SMART; SM00204; TGFB; 1.
DR   PROSITE; PS00250; TGF_BETA_1; 1.
KW   Glycoprotein.
FT   NON_TER       1       1
SQ   SEQUENCE   337 AA;  38792 MW;  BC91231431FDCA27 CRC64;

  Query Match             24.9%;  Score 341;  DB 6;  Length 337;
  Best Local Similarity   35.4%;  Pred. No. 6e-24;
  Matches   84;  Conservative   29;  Mismatches   76;  Indels   48;  Gaps    7;

Qy      29 WETLEVSSAVKRWVRADSTTNKNKLEVT----VQSHRESCDTLDISVPPGSKN----LPF 80
           ||| :|| || || |          :|||     ::|:     :  |:| || :     |
Db     141 WETFDVSPAVLRWTREKQPNYGLAIEVTHLHQTRTHQGQHVRISRSLPQGSGDWAQLRPL 200

Qy      81 FVVFSNDRSNGTKETRLDLLKEMIGHEQETMLVKTAKNAYQGAGESQEEEGLDGYTAVGP 140
             | | :|    |   ||     |    | :
Db     201 LVTFGHD-GRGHALTRRQRAKRSPKHHAQR------------------------------ 229

Qy     141 LLARRKRSTGASSHCQKTSLRVNFEDIGWDSWIIAPKEYDAYECKGGCFFPLADDVTPTK 200
             ||:|     : :|:: || |:| |:||: ||:||  | |: | | | |||||: :  |
Db     230 --ARKK-----NKNCRRHSLYVDFSDVGWNDWIVAPPGYQAFYCHGDCPFPLADHLNSTN 282

Qy     201 HAIVQTLVHLKFPTKVGKACCVPTKLSPISILYKDDMGVPTLKYHYEGMSVAECGCR 257
           ||||||||:    : : |||||||:|| ||:|| |:     || :|: | |  ||||
Db     283 HAIVQTLVN-SVNSSIPKACCVPTELSAISMLYLDEYDKVVLK-NYQEMVVEGCGCR 337


RESULT 5
Q9W753
ID   Q9W753      PRELIMINARY;      PRT;   399 AA.
AC   Q9W753;
DT   01-NOV-1999 (TrEMBLrel. 12, Created)
DT   01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT   01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE   Growth and differentiation factor 6.
GN   GDF6.
OS   Xenopus laevis (African clawed frog).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidea; Pipidae;
OC   Xenopodinae; Xenopus.
OX   NCBI_TaxID=8355;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=99396700; PubMed=10393114;
RA   Chang C., Hemmati-Brivanlou A.;
RT   "Xenopus GDF6, a new antagonist of noggin and a partner of BMPs.";
RL   Development 126:3347-3357(1999).
CC   -!- SIMILARITY: BELONGS TO THE TGF-BETA FAMILY.
DR   EMBL; AF155125; AAD38402.1; -.
DR   HSSP; P12643; 3BMP.
DR   InterPro; IPR002405; Inhibin_alpha.
DR   InterPro; IPR001839; TGFb.
DR   InterPro; IPR001111; TGFb_N.
DR   Pfam; PF00019; TGF-beta; 1.
DR   Pfam; PF00688; TGFb_propeptide; 1.
DR   PRINTS; PR00669; INHIBINA.
DR   ProDom; PD000357; TGFb; 1.
DR   SMART; SM00204; TGFB; 1.
DR   PROSITE; PS00250; TGF_BETA_1; 1.
KW   Glycoprotein.
SQ   SEQUENCE   399 AA;  45571 MW;  C549D973B50B8517 CRC64;

  Query Match             24.5%;  Score 335.5;  DB 13;  Length 399;
  Best Local Similarity   34.2%;  Pred. No. 2.5e-23;
  Matches   89;  Conservative   34;  Mismatches   76;  Indels   61;  Gaps   11;

Qy      23 DIRD---EGWETLEVSSAVKRWVRADSTTNKNKLEVTVQSHRESCDTLD-ISVPPGSKNL 78
           |:||    ||:  :|      |           |:   |  :: |  |  ||:  | :
Db     176 DLRDTPSPGWQVFDV------W---------KSLQDISQGKKQICVELKAISLTTGLE-- 218

Qy      79 PFFVVFSNDRSNGTKETRLDLLKEMIGHEQETMLV----KTAKNAYQ------GAGESQE 128
             | ||       | | ::   |::: :||     : || |        : :| |
Db     219 ------VNLRS-------LGLARKPRSHQEKALLVVFTKSSRKNLYNELKEQVHSSKSME 265

Qy     129 EEGLDGYTAVGPLLARRKRSTGASSH----------CQKTSLRVNFEDIGWDSWIIAPK 177
           :|    :        ||:|:|  | |           | |  | |||:::||| |||||
Db     266 KEARLHFKT-----RRRRRTTFNSRHGKRHGRKSRLRCSKKPLHVNFKELGWDDWIIAPL 320

Qy     178 EYDAYECKGGCFFPLADDVTPTKHAIVQTLVHLKFPTKVGKACCVPTKLSPISILYKDDM 237
           ||:|: |:| | |||   : || |||:|||::   |     :|||||||:||||||  |
Db     321 EYEAHHCEGVCDFPLRSHLEPTNHAIIQTLMNSMNPGSTPPSCCVPTKLTPISILY-IDA 379

Qy     238 GVPTLKYHYEGMSVAECGCR 257
           |   :   || | |  ||||
Db     380 GNNVVYKQYEDMVVESCGCR 399


RESULT 6
P91720
ID   P91720      PRELIMINARY;      PRT;   614 AA.
AC   P91720;
DT   01-MAY-1997 (TrEMBLrel. 03, Created)
DT   01-MAY-1997 (TrEMBLrel. 03, Last sequence update)
DT   01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE   Decapentaplegic protein.
GN   DPP.
OS   Drosophila virilis (Fruit fly).
OC   Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC   Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC   Ephydroidea; Drosophilidae; Drosophila.
OX   NCBI_TaxID=7244;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=97225212; PubMed=9071585;
RA   Newfeld S.J., Padgett R.W., Findley S.D., Richter B.G., Sanicola M.,
RA   de Cuevas M., Gelbart W.M.;
RT   "Molecular evolution at the decapentaplegic locus in Drosophila.";
RL   Genetics 145:297-309(1997).
CC   -!- SIMILARITY: BELONGS TO THE TGF-BETA FAMILY.
DR   EMBL; U63855; AAC47555.1; -.
DR   HSSP; P12643; 3BMP.
DR   FlyBase; FBgn0013109; Dvir\dpp.
DR   InterPro; IPR001839; TGFb.
DR   InterPro; IPR001111; TGFb_N.
DR   Pfam; PF00019; TGF-beta; 1.
DR   Pfam; PF00688; TGFb_propeptide; 1.
DR   ProDom; PD000357; TGFb; 1.
DR   SMART; SM00204; TGFB; 1.
DR   PROSITE; PS00250; TGF_BETA_1; 1.
KW   Glycoprotein.
```

```
SQ   SEQUENCE   614 AA;  69055 MW;  D0F20A4093403DCF CRC64;

  Query Match           24.3%;  Score 333;  DB 5;  Length 614;
  Best Local Similarity   33.3%;  Pred. No. 7.9e-23;
  Matches   84;  Conservative   37;  Mismatches   87;  Indels   44;  Gaps    8;

Qy      19 LVSQDIRDEGWET--LEVSSAVKRWVRADSTTNKN----------KLEVTVQSHRESCD 65
           | :: :|    ||  |:|  || ||:   :|  ||           |:    |
Db     394 LDTKTVRLNSTETVSLDVQPAVDRWL---ATPQKNYGLLVEVRTMRSLKPAPHHHVRLRR 450

Qy      66 TLDISVPPGSKNLPFFVVFSNDRSNGTKETRLDLLKEMIGHEQETMLVKTAKNAYQGAGE 125
           : | :        |    :::|   |  ::|   :::: | |          | |
Db     451 SADEAHEDWQHKQPLLFTYTDD---GRHKSR--SIRDVSGRE---------GGGNGGGGR 496

Qy     126 SQEEEGLDGYTAVGPLLARRKRSTGASSHCQKTSLRVNFEDIGWDSWIIAPKEYDAYECK 185
           ::  :             || |      :|:: || |:|:|:||  ||:||  |||| |
Db     497 NRRHQ------------RRSRRKNNEDNCRRHSLYVDFQDVGWSDWIVAPPGYDAYYCH 543

Qy     186 GGCFFPLADDVTPTKHAIVQTLVHLKFPTKVGKACCVPTKLSPISILYKDDMGVPTLKYH 245
           | | ||||| :  | ||:|||||:   | || |||||||:|  ||:|| :|    || :
Db     544 GKCQFPLADHLNSTNHAVVQTLVNNLNPGKVPKACCVPTQLEGISMLYLNDQRTVVLK-N 602

Qy     246 YEGMSVAECGCR 257
           |: |:|  ||||
Db     603 YQDMTVVGCGCR 614


RESULT 7
Q9W6C0
ID   Q9W6C0      PRELIMINARY;      PRT;   261 AA.
AC   Q9W6C0;
DT   01-NOV-1999 (TrEMBLrel. 12, Created)
DT   01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT   01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE   Growth/differentiation factor 7 (Fragment).
GN   GDF7.
OS   Brachydanio rerio (Zebrafish) (Zebra danio).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC   Cyprinidae; Danio.
OX   NCBI_TaxID=7955;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=99148135; PubMed=10022976;
RA   Davidson A.J., Postlethwait J.H., Yan Y.L., Beier D.R., van Doren C.,
RA   Foernzler D., Celeste A.J., Crosier K.E., Crosier P.S.;
RT   "Isolation of zebrafish gdf7 and comparative genetic mapping of genes
RT   belonging to the Growth/Differentiation factor 5, 6, 7 subgroup of the
RT   TGF-beta superfamily.";
RL   Genome Res. 9:121-129(1999).
CC   -!- SIMILARITY: BELONGS TO THE TGF-BETA FAMILY.
DR   EMBL; AF113023; AAD20829.1; -.
DR   HSSP; P12643; 3BMP.
DR   ZFIN; ZDB-GENE-990714-1; gdf7.
DR   InterPro; IPR002400; GF_cysknot.
DR   InterPro; IPR002405; Inhibin_alpha.
DR   InterPro; IPR001839; TGFb.
DR   Pfam; PF00019; TGF-beta; 1.
DR   PRINTS; PR00438; GFCYSKNOT.
DR   PRINTS; PR00669; INHIBINA.
DR   ProDom; PD000357; TGFb; 1.
DR   SMART; SM00204; TGFB; 1.
DR   PROSITE; PS00250; TGF_BETA_1; 1.
KW   Glycoprotein.
FT   NON_TER       1       1
SQ   SEQUENCE   261 AA;  29414 MW;  77346E977036A104 CRC64;

  Query Match           24.1%;  Score 329.5;  DB 13;  Length 261;
  Best Local Similarity   31.6%;  Pred. No. 5.1e-23;
  Matches   84;  Conservative   32;  Mismatches   95;  Indels   55;  Gaps    7;

Qy      23 DIRDEGWETLEVSSAVKRWVRADSTTNKNK---LEVTVQSHRESCDTLDISVPPGSKNL- 78
           |    |:  :|   :|  ::   |    :   | ::  |     |: : :| ||   |

Db      20 DTSSATWDVFDVGPIIKTPLKQHRTAEDTRLLCLSISAVS-----DSNNEAVHPGMLGLS 74

Qy      79 ---------PFFVVFSNDRSNGTKETRLDLLKEMI-----------------GHEQETM 111
                     | ||  |    ||     ::| |                    || :
Db      75 REDQQTHERALLVAFSQAR---RKENLFREIREKIRAMKSRKFSNPTPEHSIKGHPRHRR 131

Qy     112 LVKTAKNAYQGAGESQEEEGLDGYTAVGPLLARRKRSTGASSHCQKTSLRVNFEDIGWDS 171
             :||                | |   |||: :  |      : | :  | |||:::|||
Db     132 RRRTA--------------LAGRPGVGPITSGGKGGGRRRTRCSRKPLHVNFKELGWDD 176

Qy     172 WIIAPKEYDAYECKGGCFFPLADDVTPTKHAIVQTLVHLKFPTKVGKACCVPTKLSPISI 231
           ||||| :|:|| |:| | |||   : || |||:|||::   |     :||||:|||||||
Db     177 WIIAPLDYEAYHCEGLCDFPLRSHLEPTNHAIIQTLMNSMDPESTPPSCCVPSKLSPISI 236

Qy     232 LYKDDMGVPTLKYHYEGMSVAECGCR 257
           || |  |   :   || | |  ||||
Db     237 LYIDS-GNNVVYKQYEDMVVESCGCR 261


RESULT 8
Q8UVQ8
ID   Q8UVQ8      PRELIMINARY;      PRT;   391 AA.
AC   Q8UVQ8;
DT   01-MAR-2002 (TrEMBLrel. 20, Created)
DT   01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT   01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE   Anti-dorsalizing morphogenetic protein.
OS   Brachydanio rerio (Zebrafish) (Zebra danio).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC   Cyprinidae; Danio.
OX   NCBI_TaxID=7955;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=21614694; PubMed=11748836;
RA   Lele Z., Nowak M., Hammerschmidt M.;
RT   "Zebrafish admp is required to restrict the size of the organizer and
RT   to promote posterior and ventral development.";
RL   Dev. Dyn. 222:681-687(2001).
DR   EMBL; AF418564; AAL49502.1; -.
DR   InterPro; IPR001839; TGFb.
DR   InterPro; IPR001111; TGFb_N.
DR   Pfam; PF00019; TGF-beta; 1.
DR   Pfam; PF00688; TGFb_propeptide; 1.
DR   ProDom; PD000357; TGFb; 1.
DR   SMART; SM00204; TGFB; 1.
DR   PROSITE; PS00250; TGF_BETA_1; UNKNOWN_1.
SQ   SEQUENCE   391 AA;  43792 MW;  F89885D22364962B CRC64;

  Query Match           23.9%;  Score 326.5;  DB 13;  Length 391;
  Best Local Similarity   32.3%;  Pred. No. 1.7e-22;
  Matches   87;  Conservative   43;  Mismatches   92;  Indels   47;  Gaps   11;

Qy       5 DSETWDQATGTKTFLVSQ---DIRDEGWETLEVSSAVKRWVRADSTTNKNKLEVTVQSHR 61
           ||   : : | |. |:|     |   |||   :: ||: |: :|  :|   | |:|::
Db     154 DSSKKNVSQGKK--LLSSRLVPIHSTGWEVFTITQAVRSWM-SDEGSNLGLL-VSVRTLA 209

Qy      62 ESCDTLDISVPPGSKNLPFFVVFSNDRSNGTKETRLDLLKEMIGHEQETMLVKTAKNAYQ 121
           |   :|: :          | |:: | :                | :: |||    :  :
Db     210 GS--QMDLKM----------VRFASGRDH-------------HHSKQPMLVLFTDDGRR 243

Qy     122 GAGESQEEEGLD----GYTAVGPLLARRKRST--------GASSHCQKTSLRVNFEDIGW 169
           |      :| |    | :   | :   :||:         |    ||:  | |:||:|||
Db     244 AASLEATSKGSDVSPGGXSQPLPSVPASRRSSRSVDYDERGEKMACQRQPLYVDFEEIGW 303

Qy     170 DSWIIAPKEYDAYECKGGCFFPLADDVTPTKHAIVQTLVH-LKFPTKVGKACCVPTKLSP 228
           ||::|| |:|| ||| | |||: :: || ||||:::: ||    :   |||| ||
Db     304 SGWIVSPKGYNAYHCKGSCIFPLSQNMRPTNHAIVQSIINTLKLNKGIQTPCCVPDKLYS 363

Qy     229 ISILYKDDMGVPTLKYHYEGMSVAECGCR 257
           ||:|| ||    ||  |: |    ||||
Db     364 ISLLYFDDDENVVLK-QYDDMVAGSCGCR 391
```

```
RESULT 9
O73818
ID   O73818      PRELIMINARY;      PRT;   400 AA.
AC   O73818;
DT   01-AUG-1998 (TrEMBLrel. 07, Created)
DT   01-AUG-1998 (TrEMBLrel. 07, Last sequence update)
DT   01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE   Bone morphogenetic protein 4.
GN   BMP-4.
OS   Xenopus laevis (African clawed frog).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidea; Pipidae;
OC   Xenopodinae; Xenopus.
OX   NCBI_TaxID=8355;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=98440832; PubMed=9753664;
RA   Kim J., Ault K.T., Chen H.D., Xu R.H., Roh D.H., Lin M.C., Park M.J.,
RA   Kung H.F.;
RT   "Transcriptional regulation of BMP-4 in the Xenopus embryo: analysis
RT   of genomic BMP-4 and its promoter.";
RL   Biochem. Biophys. Res. Commun. 250:516-530(1998).
CC   -!- SIMILARITY: BELONGS TO THE TGF-BETA FAMILY.
DR   EMBL; AF058764; AAC61694.1; -.
DR   HSSP; P12643; 3BMP.
DR   InterPro; IPR001839; TGFb.
DR   InterPro; IPR001111; TGFb_N.
DR   Pfam; PF00019; TGF-beta; 1.
DR   Pfam; PF00688; TGFb_propeptide; 1.
DR   ProDom; PD000357; TGFb; 1.
DR   SMART; SM00204; TGFB; 1.
DR   PROSITE; PS00250; TGF_BETA_1; 1.
KW   Glycoprotein.
SQ   SEQUENCE   400 AA;  45810 MW;  A3147E4FACB4553F CRC64;

  Query Match             23.9%;  Score 326.5;  DB 13;  Length 400;
  Best Local Similarity   32.8%;  Pred. No. 1.8e-22;
  Matches   78;  Conservative   35;  Mismatches   72;  Indels   53;  Gaps    7;

Qy      29 WETLEVSSAVKRWVRADSTTNKNKLEV----TVQSHRESCDTLDISVPP-----GSKNLP 79
           ||: :|| |: || |     :   :||      ::|:     :  |: |      |:  |
Db     207 WESFDVSPAIMRWTRDKQINHGLAIEVIHLNQTKTHQGKHVRISRSLLPQEDADWSQMRP 266

Qy      80 FFVVFSNDRSNGTKETRLDLLKEMIGHEQETMLVKTAKNAYQGAGESQEEEGLDGYTAVG 139
             : ||:|                ||     | : :| :
Db     267 LLITFSHDGR--------------GH----ALTRRSKRS---------------------- 287

Qy     140 PLLARRKRSTGASSHCQKTSLRVNFEDIGWDSWIIAPKEYDAYECKGGCFFPLADDVTPT 199
             :::|    : ||:: || |:| |:||: ||:||  | |: | | | ||||| :  |
Db     288 ---PKQQRPRKKNKHCRRHSLYVDFSDVGWNDWIVAPPGYQAFYCHGDCPFPLADHLNST 344

Qy     200 KHAIVQTLVHLKFPTKVGKACCVPTKLSPISILYKDDMGVPTLKYHYEGMSVAECGCR 257
           ||||||||:      : |||||||:|| ||:|| |:     || :|: | |  ||||
Db     345 NHAIVQTLVN-SVNASIPKACCVPTELSAISMLYLDEYDKVVLK-NYQEMVVEGCGCR 400


RESULT 10
Q91703
ID   Q91703      PRELIMINARY;      PRT;   400 AA.
AC   Q91703;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE   Protein 4.
GN   BMP-4.
OS   Xenopus laevis (African clawed frog).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidea; Pipidae;
OC   Xenopodinae; Xenopus.
OX   NCBI_TaxID=8355;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Smith J.C.;
RL   Submitted (FEB-1992) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE OF 1-20 FROM N.A.
RX   MEDLINE=93048819; PubMed=1425340;
RA   Dale L., Howes G., Price B.M., Smith J.C.;
RT   "Bone morphogenetic protein 4: a ventralizing factor in early Xenopus
RT   development.";
RL   Development 115:573-585(1992).
RN   [3]
RP   SEQUENCE FROM N.A.
RA   Metz A., Knoechel S., Buechler P., Koester M., Knoechel W.;
RL   Submitted (MAR-1998) to the EMBL/GenBank/DDBJ databases.
CC   -!- SIMILARITY: BELONGS TO THE TGF-BETA FAMILY.
DR   EMBL; X64538; CAA45836.1; -.
DR   EMBL; AJ005076; CAA06333.1; -.
DR   HSSP; P12643; 3BMP.
DR   InterPro; IPR001839; TGFb.
DR   InterPro; IPR001111; TGFb_N.
DR   Pfam; PF00019; TGF-beta; 1.
DR   Pfam; PF00688; TGFb_propeptide; 1.
DR   ProDom; PD000357; TGFb; 1.
DR   SMART; SM00204; TGFB; 1.
DR   PROSITE; PS00250; TGF_BETA_1; 1.
KW   Glycoprotein.
SQ   SEQUENCE   400 AA;  45778 MW;  B81472F9BCB4506E CRC64;

  Query Match             23.9%;  Score 326.5;  DB 13;  Length 400;
  Best Local Similarity   32.8%;  Pred. No. 1.8e-22;
  Matches   78;  Conservative   35;  Mismatches   72;  Indels   53;  Gaps    7;

Qy      29 WETLEVSSAVKRWVRADSTTNKNKLEV----TVQSHRESCDTLDISVPP-----GSKNLP 79
           ||: :|| |: || |     :   :||      ::|:     :  |: |      |:  |
Db     207 WESFDVSPAIMRWTRDKQINHGLAIEVIHLNQTKTHQGKHVRISRSLLPQEDADWSQMRP 266

Qy      80 FFVVFSNDRSNGTKETRLDLLKEMIGHEQETMLVKTAKNAYQGAGESQEEEGLDGYTAVG 139
             : ||:|                ||     | : :| :
Db     267 LLITFSHDGR--------------GH----ALTRRSKRS---------------------- 287

Qy     140 PLLARRKRSTGASSHCQKTSLRVNFEDIGWDSWIIAPKEYDAYECKGGCFFPLADDVTPT 199
             :::|    : ||:: || |:| |:||: ||:||  | |: | | | ||||| :  |
Db     288 ---PKQQRPRKKNKHCRRHSLYVDFSDVGWNDWIVAPPGYQAFYCHGDCPFPLADHLNST 344

Qy     200 KHAIVQTLVHLKFPTKVGKACCVPTKLSPISILYKDDMGVPTLKYHYEGMSVAECGCR 257
           ||||||||:      : |||||||:|| ||:|| |:     || :|: | |  ||||
Db     345 NHAIVQTLVN-SVNASIPKACCVPTELSAISMLYLDEYDKVVLK-NYQEMVVEGCGCR 400


RESULT 11
Q8WS99
ID   Q8WS99      PRELIMINARY;      PRT;   509 AA.
AC   Q8WS99;
DT   01-MAR-2002 (TrEMBLrel. 20, Created)
DT   01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT   01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE   Bone morphogenetic protein BMP2/4.
OS   Archaster typicus.
OC   Eukaryota; Metazoa; Echinodermata; Eleutherozoa; Asterozoa;
OC   Asteroidea; Valvatacea; Valvatida; Archasteridae; Archaster.
OX   NCBI_TaxID=136937;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Shih L.-J., Chen C.-P., Hwang S.-P.L.;
RT   "Uniform Distribution of Sea Star BMP2/4 mRNA in Embryos at the Later
RT   Stages of Embryonic Development.";
RL   Submitted (JAN-2001) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF334705; AAL73188.1; -.
DR   InterPro; IPR001839; TGFb.
DR   InterPro; IPR001111; TGFb_N.
DR   Pfam; PF00019; TGF-beta; 1.
DR   Pfam; PF00688; TGFb_propeptide; 1.
DR   ProDom; PD000357; TGFb; 1.
DR   SMART; SM00204; TGFB; 1.
```

```
DR   PROSITE; PS00250; TGF_BETA_1; UNKNOWN_1.
SQ   SEQUENCE   509 AA;  57015 MW;  EC4D40732A33B735 CRC64;

  Query Match           23.9%;  Score 326.5;  DB 5;  Length 509;
  Best Local Similarity  31.6%;  Pred. No. 2.5e-22;
  Matches   80;  Conservative   33;  Mismatches   77;  Indels   63;  Gaps    6;

Qy      23 DIRDEGWETLEVSSAVKRWVRADSTTNKNKLEV------TVQSHRESCDTLDISVPPGS- 75
           |||:  | : :|  |:: | |:    :  ::||       : :||       :    |:
Db     302 DIRNVSWHSFDVRPAIRDWRRSHLANHGIEVEVLDHRGRPIAAHRHLRIARSVGGEYGNA 361

Qy      76 -------KNLPFFVVFSND----RSNGTKETRLDLLKEMIGHEQETMLVKTAKNAYQGAG 124
                  :  |  | |::|    ||  || :
Db     362 ADETRWFEERPLIVTFTDDGRRKRSATTKRS---------------------------- 392

Qy     125 ESQEEEGLDGYTAVGPLLARRKRSTGASSHCQKTSLRVNFEDIGWDSWIIAPKEYDAYEC 184
            |::  |              ||      :|::  | |:| |:||:|||:||  | || |
Db     393 -SRQRSG-------------KRKRKLKPNCRRHPLYVDFTDVGWNSWIVAPAGYQAYYC 437

Qy     185 KGGCFFPLADDVTPTKHAIVQTLVHLKFPTKVGKACCVPTKLSPISILYKDDMGVPTLKY 244
           :| | ||| | :  | ||||||||:   |    |||||| || ||:|| ||     |:
Db     438 QGECPFPLVDHLNATNHAIVQTLVNSASPQLAPKACCVPTDLSAISMLYLDDSDSVILR- 496

Qy     245 HYEGMSVAECGCR 257
           :|: | |  ||||
Db     497 NYQDMVVEGCGCR 509


RESULT 12
Q8UVQ2
ID   Q8UVQ2      PRELIMINARY;      PRT;   391 AA.
AC   Q8UVQ2;
DT   01-MAR-2002 (TrEMBLrel. 20, Created)
DT   01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT   01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE   Anti-dorsalizing morphogenetic protein.
OS   Brachydanio rerio (Zebrafish) (Zebra danio).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC   Cyprinidae; Danio.
OX   NCBI_TaxID=7955;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=21850182; PubMed=11862464;
RA   Dickmeis T., Rastegar S., Aanstad P., Clark M., Fischer N., Korzh V.,
RA   Straehle U.;
RT   "Expression of the anti-dorsalizing morphogenetic protein gene in the
RT   zebrafish embryo.";
RL   Dev. Genes Evol. 211:568-572(2001).
DR   EMBL; AF420475; AAL60179.1; -.
DR   InterPro; IPR001839; TGFb.
DR   InterPro; IPR001111; TGFb_N.
DR   Pfam; PF00019; TGF-beta; 1.
DR   Pfam; PF00688; TGFb_propeptide; 1.
DR   ProDom; PD000357; TGFb; 1.
DR   SMART; SM00204; TGFB; 1.
DR   PROSITE; PS00250; TGF_BETA_1; UNKNOWN_1.
SQ   SEQUENCE   391 AA;  43769 MW;  F17F83652C34962B CRC64;

  Query Match           23.8%;  Score 325.5;  DB 13;  Length 391;
  Best Local Similarity  32.3%;  Pred. No. 2.2e-22;
  Matches   87;  Conservative   43;  Mismatches   92;  Indels   47;  Gaps   11;

Qy       5 DSETWDQATGTKTFLVSQ---DIRDEGWETLEVSSAVKRWVRADSTTNKNKLEVTVQSHR 61
           ||   : : | |  |:|     |   |||   :: ||: |: :|  :|   | |:|::
Db     154 DSSKKNVSQGKK--LLSSRLVPIHSTGWEVFTITQAVRSWM-SDEGSNLGLL-VSVRTLA 209

Qy      62 ESCDTLDISVPPGSKNLPFFVVFSNDRSNGTKETRLDLLKEMIGHEQETMLVKTAKNAYQ 121
             |   :|: :          | |:: | :               | :: |||    :  :
Db     210 GS--QMDLKM----------VRFASGRDH-------------HHSKQPMLVLFTDDGRR 243

Qy     122 GAGESQEEEGLD----GYTAVGPLLARRKRST--------GASSHCQKTSLRVNFEDIGW 169
             |      :| |    | :   | :   :||:        |    ||:  | |:||:|||
```

```
Db     244 AASLEATSKGSDVSPGGSSQPLPSVPASRRSSRSVDYDERGEKMACQRQPLYVDFEEIGW 303

Qy     170 DSWIIAPKEYDAYECKGGCFFPLADDVTPTKHAIVQTLVH-LKFPTKVGKACCVPTKLSP 228
           ||::|| |:|| ||| | |||: :: || |||||:::: ||    :   |||| ||
Db     304 SGWIVSPKGYNAYHCKGSCIFPLSQNMRPTNHAIVQSIINTLKLNKGIQTPCCVPDKLYS 363

Qy     229 ISILYKDDMGVPTLKYHYEGMSVAECGCR 257
           ||:|| ||     ||  |: |    ||||
Db     364 ISLLYFDDDENVVLK-QYDDMVAGSCGCR 391


RESULT 13
Q9VQC6
ID   Q9VQC6      PRELIMINARY;      PRT;   588 AA.
AC   Q9VQC6;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE   DPP protein.
GN   DPP OR CG9885.
OS   Drosophila melanogaster (Fruit fly).
OC   Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC   Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC   Ephydroidea; Drosophilidae; Drosophila.
OX   NCBI_TaxID=7227;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=BERKELEY;
RX   MEDLINE=20196006; PubMed=10731132;
RA   Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA   Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA   George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA   Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA   Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA   Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA   Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA   Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA   Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA   Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
RA   Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA   Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA   de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA   Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA   Durbin K.J., Evangelista C.C., Ferraz C., Ferriera S., Fleischmann W.,
RA   Fosler C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA   Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA   Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA   Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA   Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA   Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA   Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA   Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA   Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA   Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA   Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
RA   Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA   Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA   Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA   Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA   Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA   Wang Z.-Y., Wassarman D.A., Weinstock G.M., Weissenbach J.,
RA   Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA   Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA   Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA   Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT   "The genome sequence of Drosophila melanogaster.";
RL   Science 287:2185-2195(2000).
CC   -!- SIMILARITY: BELONGS TO THE TGF-BETA FAMILY.
DR   EMBL; AE003583; AAF51250.1; -.
DR   HSSP; P12643; 3BMP.
DR   FlyBase; FBgn0000490; dpp.
DR   InterPro; IPR002405; Inhibin_alpha.
DR   InterPro; IPR001839; TGFb.
```

```
DR   InterPro; IPR001111; TGFb_N.
DR   Pfam; PF00019; TGF-beta; 1.
DR   Pfam; PF00688; TGFb_propeptide; 1.
DR   PRINTS; PR00669; INHIBINA.
DR   ProDom; PD000357; TGFb; 1.
DR   SMART; SM00204; TGFB; 1.
DR   PROSITE; PS00250; TGF_BETA_1; 1.
KW   Glycoprotein.
SQ   SEQUENCE   588 AA;  65867 MW;  2C8166C1BD2F666B CRC64;

  Query Match             23.8%;  Score 325;  DB 5;  Length 588;
  Best Local Similarity   30.5%;  Pred. No. 4.2e-22;
  Matches   92;  Conservative   46;  Mismatches  114;  Indels   50;  Gaps    9;

Qy       4 EDSETWDQATGTKTFLVSQDIR----DEGWETLEV----SSAVKRWVRADSTTNKNKLEV 55
           :||:  |:       | :  |::    ||  :  |:     :  :: | : |: |: : :|
Db     289 KDSKIDDRFPHHHRFRLHFDVKSIPADEKLKAAELQLTRDALSQQVVASRSSANRTRYQV 348

Qy      56 T--------VQSHRE-------------SCDTLDISVPPG--------SKNLPFFVVFS 85
                    |:  ||              | ||: : | |          :|    |
Db     349 LVYDITRVGVRGQREPSYLLLDTKTVRLNSTDTVSLDVQPAVDRWLASPQRNYGLLVEVR 408

Qy      86 NDRS-------NGTKETRLDLLKEMIGHEQETMLVKTAKNAYQGAGESQEEEGLDGYTAV 138
             ||       :       |   |   |:|  :   |    :: |   ::  | :|
Db     409 TVRSLKPAPHHHVRLRRSADEAHERWQHKQPLLFTYTDDGRHK-ARSIRDVSGGEGGGKG 467

Qy     139 GPLLARRKRSTGASSH---CQKTSLRVNFEDIGWDSWIIAPKEYDAYECKGGCFFPLADD 195
           |    : :| |   :|    |:: || |:| |:||| ||:||  |||| | | | |||||
Db     468 GRNKRQPRRPTRRKNHDDTCRRHSLYVDFSDVGWDDWIVAPLGYDAYYCHGKCPFPLADH 527

Qy     196 VTPTKHAIVQTLVHLKFPTKVGKACCVPTKLSPISILYKDDMGVPTLKYHYEGMSVAECG 255
             | ||:|||||:   | || |||||||:|  :::|| :|     || :|: |:|  ||
Db     528 FNSTNHAVVQTLVNNMNPGKVPKACCVPTQLDSVAMLYLNDQSTVVLK-NYQEMTVVGCG 586

Qy     256 CR 257
           ||
Db     587 CR 588
```

```
RESULT 14
Q9GT26
ID   Q9GT26      PRELIMINARY;      PRT;   438 AA.
AC   Q9GT26;
DT   01-MAR-2001 (TrEMBLrel. 16, Created)
DT   01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT   01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE   Gbb-60a-like protein As60a.
OS   Anopheles stephensi (Indo-Pakistan malaria mosquito).
OC   Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC   Pterygota; Neoptera; Endopterygota; Diptera; Nematocera; Culicoidea;
OC   Anopheles.
OX   NCBI_TaxID=30069;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=21066562; PubMed=11145845;
RA   Crampton A.L., Luckhart S.;
RT   "Isolation and characterization of As60A, a transforming growth
RT   factor-beta gene, from the malaria vector Anopheles stephensi.";
RL   Cytokine 13:65-74(2001).
CC   -!- SIMILARITY: BELONGS TO THE TGF-BETA FAMILY.
DR   EMBL; AF284816; AAG13400.1; -.
DR   HSSP; P18075; 1BMP.
DR   InterPro; IPR001839; TGFb.
DR   InterPro; IPR001111; TGFb_N.
DR   Pfam; PF00019; TGF-beta; 1.
DR   Pfam; PF00688; TGFb_propeptide; 1.
DR   ProDom; PD000357; TGFb; 1.
DR   SMART; SM00204; TGFB; 1.
DR   PROSITE; PS00250; TGF_BETA_1; 1.
KW   Glycoprotein.
SQ   SEQUENCE   438 AA;  49824 MW;  124BA66DAA832E84 CRC64;

  Query Match             23.4%;  Score 320.5;  DB 5;  Length 438;
  Best Local Similarity   30.4%;  Pred. No. 7.5e-22;
  Matches   82;  Conservative   21;  Mismatches   70;  Indels   97;  Gaps    6;

Qy      27 EGWETLEVSSAVKRWVR-------------ADSTTNKNKLEVTVQSHRESCDTLDISVP 72
           |||  : |: ||  |::               |    : ||:          |   :|
Db     225 EGWLEINVTGAVNLWLKNRQANHGLYIGAYFGDRVEREVKLD----------DIGFVSAR 274

Qy      73 PGSKNLPFFVVFSNDRSNGTKETRLDLLKEMIGHEQETMLVKTAKNAYQGAGESQEEEGL 132
              :  || ||::| :    |
Db     275 GSDEYQPFLVVYANSQQQMMK---------------------------------------- 295

Qy     133 DGYTAVGPLLAR-----------RKRSTGASSH--------------CQKTSLRVNFED 166
               |:| |             |||  | : |                |:   | |:|:|
Db     296 -------PILQRHLTRNKRSQPSRKRKPGKTEHRHPFQYHQPYDQHKSCRIQQLYVSFKD 348

Qy     167 IGWDSWIIAPKEYDAYECKGGCFFPLADDVTPTKHAIVQTLVHLKFPTKVGKACCVPTKL 226
           : |  |||||: | || | | | |||    :  | ||||||||||  |||| | || ||||
Db     349 LQWHEWIIAPEGYGAYYCSGECNFPLNAHMNATNHAIVQTLVHLNHPTKVPKPCCAPTKL 408

Qy     227 SPISILYKDDMGVPTLKYHYEGMSVAECGC 256
             |||:||  |     ||  |: | |  |||
Db     409 IPISVLYHIDEANVNLK-KYKNMVVKSCGC 437
```

```
RESULT 15
Q90YJ3
ID   Q90YJ3      PRELIMINARY;      PRT;   391 AA.
AC   Q90YJ3;
DT   01-DEC-2001 (TrEMBLrel. 19, Created)
DT   01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT   01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE   Anti-dorsalizing morphogenetic protein.
GN   ADMP.
OS   Brachydanio rerio (Zebrafish) (Zebra danio).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC   Cyprinidae; Danio.
OX   NCBI_TaxID=7955;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Peyrieras N.;
RL   Submitted (JUL-2001) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE FROM N.A.
RA   Willot V.;
RT   "Cooperative action of ADMP and BMP mediated pathways in regulating
RT   cell fates in the zebrafish gastrula.";
RL   Submitted (AUG-2001) to the EMBL/GenBank/DDBJ databases.
CC   -!- SIMILARITY: BELONGS TO THE TGF-BETA FAMILY.
DR   EMBL; AJ315468; CAC50881.1; -.
DR   InterPro; IPR001839; TGFb.
DR   InterPro; IPR001111; TGFb_N.
DR   Pfam; PF00019; TGF-beta; 1.
DR   Pfam; PF00688; TGFb_propeptide; 1.
DR   ProDom; PD000357; TGFb; 1.
DR   PROSITE; PS00250; TGF_BETA_1; UNKNOWN_1.
KW   Glycoprotein.
SQ   SEQUENCE   391 AA;  43789 MW;  69CC73A98702BBD7 CRC64;

  Query Match             23.4%;  Score 320;  DB 13;  Length 391;
  Best Local Similarity   32.9%;  Pred. No. 7.1e-22;
  Matches   82;  Conservative   37;  Mismatches   84;  Indels   46;  Gaps   10;

Qy      24 IRDEGWETLEVSSAVKRWVRADSTTNKNKLEVTVQSHRESCDTLDISVPPGSKNLPFFVV 83
           |   |||   :: ||: |: :|  :|   | |:|::   |   :|: :          |
Db     174 IHSTGWEVFTITQAVRSWM-SDEGSNLGLL-VSVRTLAGS--QMDLKM----------VR 219

Qy      84 FSNDRSNGTKETRLDLLKEMIGHEQETMLVKTAKNAYQGAGESQEEEGLDGYTAVG---- 139
           |:: | :               | :: |||    :  : | |      :| |   : |
Db     220 FASGRDH--------------HHSKQPMLVLFTDDGRRAASLEATSKGSD--VSPGSPSQ 263

Qy     140 -----PLLARRKRST-----GASSHCQKTSLRVNFEDIGWDSWIIAPKEYDAYECKGGCF 189
                |   |  ||      |    ||:  | |:||:|||  ||::|| |:|| ||| |
```

```
Db    264 PLPSVPASRRSPRSVDYDERGEKMACQRQPLYVDFEEIGWSGWIVSPKGYNAYHCKGSCI 323

Qy    190 FPLADDVTPTKHAIVQTLVH-LKFPTKVGKACCVPTKLSPISILYKDDMGVPTLKYHYEG 248
          |||: :: || |||||:::: ||    :   |||| ||  ||:|| ||     ||  |:
Db    324 FPLSQNMRPTNHAIVQSIINTLKLNKGIQTPCCVPDKLYSISLLYFDDDENVVLK-QYDD 382

Qy    249 MSVAECGCR 257
          |    ||||
Db    383 MVAGSCGCR 391


Search completed: April  2, 2003, 14:51:41
Job time : 88 secs
```

GenCore version 5.1.4_p5_4578

OM protein - protein search, using sw model

Run on: April 2, 2003, 14:40:40 ; Search time 26 Seconds
(without alignments)
409.978 Million cell updates/sec

Title: US-10-002-278-9
Perfect score: 1368
Sequence: 1 DVLEDSETWDQATGTKTFLV.....GVPTLKYHYEGMSVAECGCR 257

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_40:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	1349.5	98.6	428	1	GDF2_MOUSE	Q9wv56 mus musculu
2	1146.5	83.8	429	1	GDF2_HUMAN	Q9uk05 homo sapien
3	808.5	59.1	427	1	DSL1_CHICK	P34822 gallus gall
4	516.5	37.8	424	1	BM10_HUMAN	O95393 homo sapien
5	507	37.1	420	1	BM10_MOUSE	Q9r229 mus musculu
6	345	25.2	408	1	BMP4_HUMAN	P12644 homo sapien
7	342.5	25.0	408	1	BMP4_RAT	Q06826 rattus norv
8	341	24.9	408	1	BMP4_DAMDA	Q29607 dama dama (
9	337.5	24.7	408	1	BMP4_MOUSE	P21275 mus musculu
10	336	24.6	409	1	BMP4_RABIT	O46576 oryctolagus
11	330	24.1	436	1	GDF6_BOVIN	P55106 bos taurus
12	327.5	23.9	405	1	BMP4_CHICK	Q90752 gallus gall
13	325	23.8	588	1	DECA_DROME	P07713 drosophila
14	324.5	23.7	621	1	DECA_DROPS	P91699 drosophila
15	323	23.6	593	1	DECA_DROSI	P91706 drosophila
16	319	23.3	398	1	BMPB_XENLA	P30884 xenopus lae
17	316.5	23.1	401	1	BMP4_XENLA	P30885 xenopus lae
18	316	23.1	395	1	BMP2_RABIT	O46564 oryctolagus
19	315	23.0	398	1	BMPA_XENLA	P25703 xenopus lae
20	314.5	23.0	430	1	BMP7_MOUSE	P23359 mus musculu
21	314	23.0	393	1	BMP2_RAT	P49001 rattus norv
22	313	22.9	394	1	BMP2_MOUSE	P21274 mus musculu
23	312.5	22.8	454	1	BMP5_HUMAN	P22003 homo sapien
24	312	22.8	125	1	GDF6_MOUSE	P43028 mus musculu
25	311	22.7	396	1	BMP2_HUMAN	P12643 homo sapien
26	311	22.7	399	1	BM8A_MOUSE	P34821 mus musculu
27	310.5	22.7	461	1	DVR1_STRPU	P48969 strongyloce
28	310	22.7	396	1	BMP2_DAMDA	O19006 dama dama (
29	307	22.4	353	1	BMP2_CHICK	Q90751 gallus gall
30	307	22.4	501	1	GDF5_HUMAN	P43026 homo sapien
31	305.5	22.3	431	1	BMP7_HUMAN	P18075 homo sapien
32	305.5	22.3	452	1	BMP5_MOUSE	P49003 mus musculu
33	305.5	22.3	495	1	GDF5_MOUSE	P43027 mus musculu
34	301	22.0	402	1	BMP8_HUMAN	P34820 homo sapien
35	301	22.0	510	1	BMP6_MOUSE	P20722 mus musculu
36	298	21.8	513	1	BMP6_HUMAN	P22004 homo sapien
37	297	21.7	426	1	BMP7_XENLA	P30886 xenopus lae
38	294.5	21.5	436	1	60A_DROVI	Q24735 drosophila
39	294	21.5	151	1	GDF7_MOUSE	P43029 mus musculu
40	290	21.2	360	1	DVR1_XENLA	P09534 xenopus lae
41	289.5	21.2	355	1	DVR1_BRARE	P35621 brachydanio
42	285	20.8	354	1	NODA_MOUSE	P43021 mus musculu
43	281	20.5	207	1	BMP6_RAT	Q04906 rattus norv
44	276.5	20.2	372	1	DECA_TRICA	Q26974 tribolium c
45	274.5	20.1	455	1	60A_DROME	P27091 drosophila

ALIGNMENTS

RESULT 1
GDF2_MOUSE
ID GDF2_MOUSE STANDARD; PRT; 428 AA.
AC Q9WV56; Q9QZE0;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Growth/differentiation factor 2 precursor (GDF-2) (Bone morphogenetic
DE protein 9) (BMP-9).
GN GDF2 OR BMP9.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CD-1; TISSUE=Liver;
RA Zimmers T.A., Koniaris L.G., Sitzmann J.V., Lee S.-J.;
RT "Growth/differentiation factor-2, a new TGF-beta family member with
RT bone promoting activities.";
RL Submitted (JUN-1999) to the EMBL/GenBank/DDBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Liver;
RA Celeste A.J.;
RL Submitted (SEP-1999) to the EMBL/GenBank/DDBJ databases.
CC -!- FUNCTION: COULD BE INVOLVED IN BONE FORMATION.
CC -!- SUBUNIT: HOMODIMER; DISULFIDE-LINKED (BY SIMILARITY).
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- SIMILARITY: BELONGS TO THE TGF-BETA FAMILY.
CC ----------

CC ----------
DR EMBL; AF156890; AAD40308.1; -.
DR EMBL; AF188286; AAD56961.1; -.
DR HSSP; P12643; 3BMP.
DR MGD; MGI:1321394; Gdf2.
DR InterPro; IPR002405; Inhibin_alpha.
DR InterPro; IPR001839; TGFb.
DR InterPro; IPR001111; TGFb_N.
DR Pfam; PF00019; TGF-beta; 1.
DR Pfam; PF00688; TGFb_propeptide; 1.
DR PRINTS; PR00669; INHIBINA.
DR ProDom; PD000357; TGFb; 1.
DR SMART; SM00204; TGFB; 1.
DR PROSITE; PS00250; TGF_BETA_1; 1.
KW Signal; Growth factor; Cytokine; Glycoprotein.
FT SIGNAL 1 22 POTENTIAL.
FT PROPEP 23 318 BY SIMILARITY.
FT CHAIN 319 428 GROWTH/DIFFERENTIATION FACTOR 2.

```
FT   DISULFID    326    392       BY SIMILARITY.
FT   DISULFID    355    425       BY SIMILARITY.
FT   DISULFID    359    427       BY SIMILARITY.
FT   DISULFID    391    391       INTERCHAIN (BY SIMILARITY).
FT   CARBOHYD     70     70       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    135    135       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    262    262       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CONFLICT    293    293       G -> V (IN REF. 2).
FT   CONFLICT    382    382       K -> E (IN REF. 2).
SQ   SEQUENCE   428 AA;  47660 MW;  6B81525FF8A76A39 CRC64;

  Query Match             98.6%;  Score 1349.5;  DB 1;  Length 428;
  Best Local Similarity   99.2%;  Pred. No. 1.1e-112;
  Matches  255;  Conservative    1;  Mismatches    0;  Indels    1;  Gaps    1;

Qy       1 DVLEDSETWDQATGTKTFLVSQDIRDEGWETLEVSSAVKRWVRADSTTNKNKLEVTVQSH 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     173 DVLEDSETWDQATGTKTFLVSQDIRDEGWETLEVSSAVKRWVRADSTTNKNKLEVTVQSH 232

Qy      61 RESCDTLDISVPPGSKNLPFFVVFSNDRSNGTKETRLDLLKEMIGHEQETMLVKTAKNAY 120
           ||||||||||||||||||||||||||||||||||: ||||||||||||||||||||||
Db     233 RESCDTLDISVPPGSKNLPFFVVFSNDRSNGTKETRLE-LKEMIGHEQETMLVKTAKNAY 291

Qy     121 QGAGESQEEEGLDGYTAVGPLLARRKRSTGASSHCQKTSLRVNFEDIGWDSWIIAPKEYD 180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     292 QGAGESQEEEGLDGYTAVGPLLARRKRSTGASSHCQKTSLRVNFEDIGWDSWIIAPKEYD 351

Qy     181 AYECKGGCFFPLADDVTPTKHAIVQTLVHLKFPTKVGKACCVPTKLSPISILYKDDMGVP 240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     352 AYECKGGCFFPLADDVTPTKHAIVQTLVHLKFPTKVGKACCVPTKLSPISILYKDDMGVP 411

Qy     241 TLKYHYEGMSVAECGCR 257
           |||||||||||||||||
Db     412 TLKYHYEGMSVAECGCR 428


RESULT 2
GDF2_HUMAN
ID   GDF2_HUMAN     STANDARD;      PRT;   429 AA.
AC   Q9UK05; Q9Y571;
DT   16-OCT-2001 (Rel. 40, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Growth/differentiation factor 2 precursor (GDF-2) (Bone morphogenetic
DE   protein 9) (BMP-9).
GN   GDF2 OR BMP9.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Liver;
RA   Celeste A.J.;
RL   Submitted (SEP-1999) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE OF 316-429 FROM N.A.
RC   TISSUE=Liver;
RA   Zimmers T.A., Koniaris L.G., Sitzmann J.V., Lee S.-J.;
RT   "Growth/differentiation factor-2, a new TGF-beta family member with
RT   bone promoting activities.";
RL   Submitted (JUN-1999) to the EMBL/GenBank/DDBJ databases.
CC   -!- FUNCTION: COULD BE INVOLVED IN BONE FORMATION.
CC   -!- SUBUNIT: HOMODIMER; DISULFIDE-LINKED (BY SIMILARITY).
CC   -!- SUBCELLULAR LOCATION: Secreted.
CC   -!- SIMILARITY: BELONGS TO THE TGF-BETA FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AF188285; AAD56960.1; -.
DR   EMBL; AF156891; AAD40309.1; -.
DR   HSSP; P18075; 1BMP.
DR   Genew; HGNC:4217; GDF2.
DR   MIM; 605120; -.
DR   InterPro; IPR002405; Inhibin_alpha.
DR   InterPro; IPR001839; TGFb.
DR   InterPro; IPR001111; TGFb_N.
DR   Pfam; PF00019; TGF-beta; 1.
DR   Pfam; PF00688; TGFb_propeptide; 1.
DR   PRINTS; PR00669; INHIBINA.
DR   ProDom; PD000357; TGFb; 1.
DR   SMART; SM00204; TGFB; 1.
DR   PROSITE; PS00250; TGF_BETA_1; 1.
KW   Signal; Growth factor; Cytokine; Glycoprotein.
FT   SIGNAL        1     22       POTENTIAL.
FT   PROPEP       23    319       BY SIMILARITY.
FT   CHAIN       320    429       GROWTH/DIFFERENTIATION FACTOR 2.
FT   DISULFID    327    393       BY SIMILARITY.
FT   DISULFID    356    426       BY SIMILARITY.
FT   DISULFID    360    428       BY SIMILARITY.
FT   DISULFID    392    392       INTERCHAIN (BY SIMILARITY).
FT   CARBOHYD     71     71       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    136    136       N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ   SEQUENCE   429 AA;  47320 MW;  5AC15DCA205FF086 CRC64;

  Query Match             83.8%;  Score 1146.5;  DB 1;  Length 429;
  Best Local Similarity   82.9%;  Pred. No. 1.3e-94;
  Matches  213;  Conservative   19;  Mismatches   24;  Indels    1;  Gaps    1;

Qy       1 DVLEDSETWDQATGTKTFLVSQDIRDEGWETLEVSSAVKRWVRADSTTNKNKLEVTVQSH 60
           |||: :: || || ||||||||||:||||||||||||||||:||| :|||||||:||
Db     174 DVLDGTDAWDSATETKTFLVSQDIQDEGWETLEVSSAVKRWVRSDSTKSKNKLEVTVESH 233

Qy      61 RESCDTLDISVPPGSKNLPFFVVFSNDRSNGTKETRLDLLKEMIGHEQETMLVKTAKNAY 120
           |: ||||||||||:|||||||||| |:|||||: |:||| ||||::| | :|:
Db     234 RKGCDTLDISVPPGSRNLPFFVVFSNDHSSGTKETRLE-LREMISHEQESVLKKLSKDGS 292

Qy     121 QGAGESQEEEGLDGYTAVGPLLARRKRSTGASSHCQKTSLRVNFEDIGWDSWIIAPKEYD 180
             ||||  ||  ||: | |  ||||||| || ||||||||||||||||||||||||:
Db     293 TEAGESSHEEDTDGHVAAGSTLARRKRSAGAGSHCQKTSLRVNFEDIGWDSWIIAPKEYE 352

Qy     181 AYECKGGCFFPLADDVTPTKHAIVQTLVHLKFPTKVGKACCVPTKLSPISILYKDDMGVP 240
           ||||||||||||||||||||||||||||||||||||||||||||||||:|||||||||
Db     353 AYECKGGCFFPLADDVTPTKHAIVQTLVHLKFPTKVGKACCVPTKLSPISVLYKDDMGVP 412

Qy     241 TLKYHYEGMSVAECGCR 257
           |||||||||||||||||
Db     413 TLKYHYEGMSVAECGCR 429


RESULT 3
DSL1_CHICK
ID   DSL1_CHICK     STANDARD;      PRT;   427 AA.
AC   P34822;
DT   01-FEB-1994 (Rel. 28, Created)
DT   01-FEB-1994 (Rel. 28, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Dorsalin-1 precursor (DSL-1).
GN   DSL-1.
OS   Gallus gallus (Chicken).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC   Gallus.
OX   NCBI_TaxID=9031;
RN   [1]
RP   SEQUENCE FROM N.A., AND SEQUENCE OF 319-322.
RC   TISSUE=Spinal cord;
RX   MEDLINE=93272310; PubMed=7916656;
RA   Basler K., Edlund T., Jessell T.M., Yamada T.;
RT   "Control of cell pattern in the neural tube: regulation of cell
```

```
RT   differentiation by dorsalin-1, a novel TGF beta family member.";
RL   Cell 73:687-702(1993).
CC   -!- FUNCTION: APPEARS TO REGULATE CELL DIFFERENTIATION WITHIN THE
CC       NEURAL TUBE. MAY REGULATE THE DIFFERENTIATION OF CELL TYPES ALONG
CC       THE DORSOVENTRAL AXIS OF THE NEURAL TUBE, ACTING IN CONJUNCTION
CC       WITH DISTINCT VENTRALIZING SIGNALS FROM THE NOTOCHORD AND FLOOR
CC       PLATE. CONTROLS THE CELL DIFFERENTIATION IN THE NEURAL TUBE IN
CC       SEVERAL WAYS: (1) PROMOTES THE DIFFERENTIATION OF CELL TYPES THAT
CC       DERIVE FROM THE DORSAL NEURAL TUBE. (2) ENSURES THAT THE DORSAL
CC       NEURAL TUBE IS REFRACTORY TO VENTRALIZING SPECIES FROM THE
CC       NOTOCHORD. (3) CAN DIFFUSE AND INFLUENCE THE FATE OF CELLS IN MORE
CC       VENTRAL REGIONS OF THE NEURAL TUBE.
CC   -!- SUBUNIT: HOMODIMER; DISULFIDE-LINKED.
CC   -!- SUBCELLULAR LOCATION: Secreted.
CC   -!- TISSUE SPECIFICITY: EXPRESSED SELECTIVELY IN THE DORSAL NEURAL
CC       TUBE. LOWER LEVELS SEEN IN KIDNEY AND MYOTOMAL CELLS.
CC   -!- DEVELOPMENTAL STAGE: IS NOT EXPRESSED IN NEURAL CELLS AT STAGES
CC       BEFORE NEURAL TUBE CLOSURE. IS EXPRESSED AT HIGH LEVELS IN THE
CC       DORSAL THIRD OF THE NEURAL TUBE, BEGINNING AT THE TIME OF NEURAL
CC       TUBE CLOSURE, BUT NOT BY VENTRAL NEURAL CELLS OR BY NONNEURAL
CC       CELLS. DORSAL RESTRICTION PERSISTS IN THE SPINAL CORD AT STAGES
CC       AFTER THE ONSET OF NEURONAL DIFFERENTIATION. AT LATER STAGES OF
CC       SPINAL DEVELOPMENT, IS RESTRICTED TO THE DORSOMEDIAL REGION OF
CC       THE SPINAL CORD, INCLUDING BUT NOT CONFINED TO THE ROOF PLATE.
CC   -!- SIMILARITY: BELONGS TO THE TGF-BETA FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; L12032; AAA48752.1; -.
DR   PIR; A40735; A40735.
DR   HSSP; P12643; 3BMP.
DR   InterPro; IPR001839; TGFb.
DR   InterPro; IPR001111; TGFb_N.
DR   Pfam; PF00019; TGF-beta; 1.
DR   Pfam; PF00688; TGFb_propeptide; 1.
DR   ProDom; PD000357; TGFb; 1.
DR   SMART; SM00204; TGFB; 1.
DR   PROSITE; PS00250; TGF_BETA_1; 1.
KW   Signal; Growth factor; Cytokine; Glycoprotein.
FT   SIGNAL       1     20       POTENTIAL.
FT   PROPEP      21    318
FT   CHAIN      319    427       DORSALIN-1.
FT   DISULFID   325    391       BY SIMILARITY.
FT   DISULFID   354    424       BY SIMILARITY.
FT   DISULFID   358    426       BY SIMILARITY.
FT   DISULFID   390    390       INTERCHAIN (BY SIMILARITY).
FT   CARBOHYD    71     71       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   136    136       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   265    265       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   292    292       N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ   SEQUENCE   427 AA;  48626 MW;  23AA42DC7085FABC CRC64;

  Query Match             59.1%;  Score 808.5;  DB 1;  Length 427;
  Best Local Similarity   61.9%;  Pred. No. 1.5e-64;
  Matches  161;  Conservative   31;  Mismatches   59;  Indels    9;  Gaps    5;

Qy       1 DVLEDSETWDQATGTKTFLVSQDIRDEGWETLEVSSAVKRWVRADSTTNKNKLEVTVQSH 60
             ||| | : |:    ||: ||| |:| ||| ||||||||||:||    |||||| ::|
Db     174 DVL-DGDHWENKESTKSLLVSHSIQDCGWEMFEVSSAVKRWVKADKMKTKNKLEVVIESK 232

Qy      61 RES---CDTLDISVPPGSKNLPFFVVFSNDRSNGTKETRLDLLKEMIGHEQETMLVKTAK 117
             |   |  |||:|   :|||| :|||||||||||||::: |:||| ||||::| |  |
Db     233 DLSGFPCGKLDITVTHDTKNLPLLIVFSNDRSNGTKETKVE-LREMIVHEQESVLNKLGK 291

Qy     118 NAYQGAGESQEEEGLDGYTAVGPLLARRKRSTGASSHCQKTSLRVNFEDIGWDSWIIAPK 177
           |      | :||: :          :| ||| || :||::||| |||::|||||||||||
Db     292 NDSSSEEEQREEKAI---ARPRQHSSRSKRSIGA-NHCRRTSLHVNFKEIGWDSWIIAPK 347

Qy     178 EYDAYECKGGCFFPLADDVTPTKHAIVQTLVHLKFPTKVGKACCVPTKLSPISILYKDDM 237
           :|:|:|||||||||| |:|||||||||||||||: | |  |||||||||  ||||||||
Db     348 DYEAFECKGGCFFPLTDNVTPTKHAIVQTLVHLQNPKKASKACCVPTKLDAISILYKDDA 407

Qy     238 GVPTLKYHYEGMSVAECGCR 257
           ||||| |:|||| |||||||
Db     408 GVPTLIYNYEGMKVAECGCR 427


RESULT 4
BM10_HUMAN
ID   BM10_HUMAN     STANDARD;      PRT;   424 AA.
AC   O95393;
DT   16-OCT-2001 (Rel. 40, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   15-JUN-2002 (Rel. 41, Last annotation update)
DE   Bone morphogenetic protein 10 precursor (BMP-10).
GN   BMP10.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Celeste A.J.;
RT   "Homo sapiens bone morphogenetic protein 10 (BMP-10) mRNA.";
RL   Submitted (OCT-1998) to the EMBL/GenBank/DDBJ databases.
CC   -!- SUBUNIT: HOMODIMER; DISULFIDE-LINKED (BY SIMILARITY).
CC   -!- SUBCELLULAR LOCATION: Secreted (By similarity).
CC   -!- SIMILARITY: BELONGS TO THE TGF-BETA FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AF101441; AAC77462.1; -.
DR   HSSP; P12643; 3BMP.
DR   InterPro; IPR002405; Inhibin_alpha.
DR   InterPro; IPR001839; TGFb.
DR   InterPro; IPR001111; TGFb_N.
DR   Pfam; PF00019; TGF-beta; 1.
DR   Pfam; PF00688; TGFb_propeptide; 1.
DR   PRINTS; PR00669; INHIBINA.
DR   ProDom; PD000357; TGFb; 1.
DR   SMART; SM00204; TGFB; 1.
DR   PROSITE; PS00250; TGF_BETA_1; 1.
KW   Signal; Growth factor; Cytokine; Glycoprotein.
FT   SIGNAL       1     21       POTENTIAL.
FT   PROPEP      22    316       POTENTIAL.
FT   CHAIN      317    424       BONE MORPHOGENETIC PROTEIN 10.
FT   DISULFID   323    389       BY SIMILARITY.
FT   DISULFID   352    421       BY SIMILARITY.
FT   DISULFID   356    423       BY SIMILARITY.
FT   DISULFID   388    388       INTERCHAIN (BY SIMILARITY).
FT   CARBOHYD    67     67       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   131    131       N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ   SEQUENCE   424 AA;  48047 MW;  3FDB3B7221BB2254 CRC64;

  Query Match             37.8%;  Score 516.5;  DB 1;  Length 424;
  Best Local Similarity   42.8%;  Pred. No. 1.5e-38;
  Matches  119;  Conservative   48;  Mismatches   68;  Indels   43;  Gaps   11;

Qy       1 DVLEDSETWDQATGTKTFLVSQDI--RDEGWETLEVSSAVKRWVRADSTTNKNKLEVTVQ 58
             :||| |:  ::       ||| :|   :  ||| :|: |::|| :: |:|  ::||| ::
Db     169 EVLE-SKGDNEGERNMLVLVSGEIYGTNSEWETFDVTDAIRRWQKSGSST--HQLEVHIE 225

Qy      59 S-HRESCDT----LDISVPPGSKNLPFFVVFSNDRSNGTKETRLDLLKEMIGHEQETMLV 113
           | | |: |     |:|     :|: |  :|||:|:|  : : | : | ||| |||   |
```

```
Db    226 SKHDEAEDASSGRLEIDTSAQNKHNPLLIVFSDDQS--SDKERKEELNEMISHEQLPEL- 282

Qy    114 KTAKNAYQGAGESQEEEGLDGYTAVGP----LLARRK----------RSTGASSHCQKTS 159
                  :  ||| ::: ||    ||  |           |     ::|::|
Db    283 -------------DNLGLDSFSS-GPGEEALLQMRSNIIYDSTARIRRNAKGNYCKRTP 327

Qy    160 LRVNFEDIGWDSWIIAPKEYDAYECKGGCFFPLADDVTPTKHAIVQTLVHLKFPTKVGKA 219
          | ::|::||||||||||  |:||||:| | :|||: :|||||||:| |||||   |  ||
Db    328 LYIDFKEIGWDSWIIAPPGYEAYECRGVCNYPLAEHLTPTKHAIIQALVHLKNSQKASKA 387

Qy    220 CCVPTKLSPISILYKDDMGVPTLKYHYEGMSVAECGCR 257
          ||||||| ||||||  | || | |: ||||:|:|||||
Db    388 CCVPTKLEPISILYL-DKGVVTYKFKYEGMAVSECGCR 424


RESULT 5
BM10_MOUSE
ID   BM10_MOUSE     STANDARD;      PRT;   420 AA.
AC   Q9R229; Q9Z1V8;
DT   16-OCT-2001 (Rel. 40, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Bone morphogenetic protein 10 precursor (BMP-10).
GN   BMP10.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=NIH Swiss;
RX   MEDLINE=99173786; PubMed=10072785;
RA   Neuhaus H., Rosen V., Thies R.S.;
RT   "Heart specific expression of mouse BMP-10 a novel member of the TGF-
RT   beta superfamily.";
RL   Mech. Dev. 80:181-184(1999).
RN   [2]
RP   SEQUENCE FROM N.A.
RA   Celeste A.;
RT   "Mouse bone morphogenetic protein 10 (BMP-10) genomic sequence, full
RT   coding region of exon 2.";
RL   Submitted (OCT-1998) to the EMBL/GenBank/DDBJ databases.
CC   -!- SUBUNIT: HOMODIMER; DISULFIDE-LINKED (BY SIMILARITY).
CC   -!- SUBCELLULAR LOCATION: Secreted (By similarity).
CC   -!- TISSUE SPECIFICITY: HEART-SPECIFIC.
CC   -!- SIMILARITY: BELONGS TO THE TGF-BETA FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AF101033; AAC95357.1; -.
DR   EMBL; AF101440; AAC77461.1; -.
DR   EMBL; AF101439; AAC77461.1; JOINED.
DR   HSSP; P12643; 3BMP.
DR   MGD; MGI:1338820; Bmp10.
DR   InterPro; IPR002405; Inhibin_alpha.
DR   InterPro; IPR001839; TGFb.
DR   InterPro; IPR001111; TGFb_N.
DR   Pfam; PF00019; TGF-beta; 1.
DR   Pfam; PF00688; TGFb_propeptide; 1.
DR   PRINTS; PR00669; INHIBINA.
DR   ProDom; PD000357; TGFb; 1.
DR   SMART; SM00204; TGFB; 1.
DR   PROSITE; PS00250; TGF_BETA_1; 1.
KW   Signal; Growth factor; Cytokine; Glycoprotein.
FT   SIGNAL       1     21       POTENTIAL.
FT   PROPEP      22    312       POTENTIAL.
FT   CHAIN      313    420       BONE MORPHOGENETIC PROTEIN 10.
FT   DISULFID    319    385       BY SIMILARITY.
FT   DISULFID    348    417       BY SIMILARITY.
FT   DISULFID    352    419       BY SIMILARITY.
FT   DISULFID    384    384       INTERCHAIN (BY SIMILARITY).
FT   CARBOHYD     67     67       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    130    130       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CONFLICT    111    111       E -> ED (IN REF. 2).
SQ   SEQUENCE   420 AA;  47802 MW;  724EC0D4CD4CF0B4 CRC64;

  Query Match             37.1%;  Score 507;  DB 1;  Length 420;
  Best Local Similarity   42.5%;  Pred. No. 1e-37;
  Matches  116;  Conservative   49;  Mismatches   72;  Indels   36;  Gaps   11;

Qy      1 DVLEDSETWDQATGTKTFLVSQDI--RDEGWETLEVSSAVKRWVRADSTTNKNKLEVTVQ 58
          :||| ::  ::   :   ||| :|   :  ||| :|: | :|| ::  :|  ::||: ::
Db    168 EVLESADGSEEER-SMLVLVSTEIYGTNSEWETFDVTDATRRWQKSGPST--HQLEIHIE 224

Qy     59 SHRESCD-----TLDISVPPGSKNLPFFVVFSNDRSNGTKETRLDLLKEMIGHEQETMLV 113
          | :   :      |:| :   :|: |  ||||:|:||  || : : | |:| |||:  |
Db    225 SRQNQAEDTGRGQLEIDMSAQNKHDPLLVVFSDDQSN-DKEQK-EELNELITHEQDLDLD 282

Qy    114 KTAKNAYQGAGESQEEEGL---------DGYTAVGPLLARRKRSTGASSHCQKTSLRVNF 164
            |   : |     :|| |         |  |        | |     ::|:|| | ::|
Db    283 SDA--FFSG----PDEEALLQMRSNMIDDSST--------RIRRNAKGNYCKKTPLYIDF 328

Qy    165 EDIGWDSWIIAPKEYDAYECKGGCFFPLADDVTPTKHAIVQTLVHLKFPTKVGKACCVPT 224
          ::|||||||||  |:||||:| | :|||: :|||||||:| |||||   |  |||||||
Db    329 KEIGWDSWIIAPPGYEAYECRGVCNYPLAEHLTPTKHAIIQALVHLKNSQKASKACCVPT 388

Qy    225 KLSPISILYKDDMGVPTLKYHYEGMSVAECGCR 257
          || ||||||  | || | |: ||||:|:|||||
Db    389 KLDPISILYL-DKGVVTYKFKYEGMAVSECGCR 420


RESULT 6
BMP4_HUMAN
ID   BMP4_HUMAN     STANDARD;      PRT;   408 AA.
AC   P12644; Q9UM80;
DT   01-OCT-1989 (Rel. 12, Created)
DT   01-OCT-1989 (Rel. 12, Last sequence update)
DT   15-JUN-2002 (Rel. 41, Last annotation update)
DE   Bone morphogenetic protein 4 precursor (BMP-4) (BMP-2B).
GN   BMP4 OR BMP2B OR DVR4.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=89072730; PubMed=3201241;
RA   Wozney J.M., Rosen V., Celeste A.J., Mitsock L.M., Whitters M.J.,
RA   Kriz R.W., Hewick R.M., Wang E.A.;
RT   "Novel regulators of bone formation: molecular clones and
RT   activities.";
RL   Science 242:1528-1534(1988).
RN   [2]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=98368971; PubMed=9701626;
RA   Shore E.M., Xu M., Shah P.B., Janoff H.B., Hahn G.V., Deardorff M.A.,
RA   Sovinsky L., Spinner N.B., Zasloff M.A., Wozney J.M., Kaplan F.S.;
RT   "The human bone morphogenetic protein 4 (BMP-4) gene: molecular
RT   structure and transcriptional regulation.";
RL   Calcif. Tissue Int. 63:221-229(1998).
RN   [3]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Duodenum;
RA   Strausberg R.;
RL   Submitted (JAN-2002) to the EMBL/GenBank/DDBJ databases.
RN   [4]
RP   SEQUENCE OF 7-408 FROM N.A.
RC   TISSUE=Placenta;
RX   MEDLINE=96063017; PubMed=7579580;
RA   Oida S., Iimura T., Maruoka Y., Takeda K., Sasaki S.;
```

```
RT   "Cloning and sequence of bone morphogenetic protein 4 (BMP-4) from a
RT   human placental cDNA library.";
RL   DNA Seq. 5:273-275(1995).
CC   -!- FUNCTION: INDUCES CARTILAGE AND BONE FORMATION. ALSO ACT IN
CC       MESODERM INDUCTION, TOOTH DEVELOPMENT, LIMB FORMATION AND FRACTURE
CC       REPAIR.
CC   -!- SUBUNIT: HOMODIMER; DISULFIDE-LINKED (BY SIMILARITY).
CC   -!- SUBCELLULAR LOCATION: SECRETED INTO THE EXTRACELLULAR MATRIX.
CC   -!- TISSUE SPECIFICITY: EXPRESSED IN THE LUNGS AND LOWER LEVELS SEEN
CC       IN THE KIDNEY. PRESENT ALSO IN NORMAL AND NEOPLASTIC PROSTATE
CC       TISSUES, AND PROSTATE CANCER CELL LINES.
CC   -!- SIMILARITY: BELONGS TO THE TGF-BETA FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; M22490; AAA51835.1; -.
DR   EMBL; U43842; AAC72278.1; -.
DR   EMBL; BC020546; AAH20546.1; -.
DR   EMBL; D30751; BAA06410.1; -.
DR   PIR; C37278; C37278.
DR   HSSP; P12643; 3BMP.
DR   Genew; HGNC:1071; BMP4.
DR   MIM; 112262; -.
DR   InterPro; IPR001839; TGFb.
DR   InterPro; IPR001111; TGFb_N.
DR   Pfam; PF00019; TGF-beta; 1.
DR   Pfam; PF00688; TGFb_propeptide; 1.
DR   ProDom; PD000357; TGFb; 1.
DR   SMART; SM00204; TGFB; 1.
DR   PROSITE; PS00250; TGF_BETA_1; 1.
KW   Signal; Growth factor; Cytokine; Bone; Cartilage; Glycoprotein.
FT   SIGNAL       1     19       POTENTIAL.
FT   PROPEP      20    292
FT   CHAIN      293    408       BONE MORPHOGENETIC PROTEIN 4.
FT   DISULFID    308    373       BY SIMILARITY.
FT   DISULFID    337    405       BY SIMILARITY.
FT   DISULFID    341    407       BY SIMILARITY.
FT   DISULFID    372    372       INTERCHAIN (BY SIMILARITY).
FT   CARBOHYD    143    143       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    208    208       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    350    350       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    365    365       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CONFLICT    152    152       V -> A (IN REF. 2 AND 3).
SQ   SEQUENCE   408 AA;  46555 MW;  79B01179DBB98204 CRC64;

  Query Match             25.2%;  Score 345;  DB 1;  Length 408;
  Best Local Similarity   35.9%;  Pred. No. 2.7e-23;
  Matches   85;  Conservative   28;  Mismatches   76;  Indels   48;  Gaps    7;

Qy      29 WETLEVSSAVKRWVRADSTTNKNKLEVT----VQSHRESCDTLDISVPPGSKN----LPF 80
           ||| :|| || || |         :|||      ::|:      :  |:| || |     |
Db     212 WETFDVSPAVLRWTREKQPNYGLAIEVTHLHQTRTHQGQHVRISRSLPQGSGNWAQLRPL 271

Qy      81 FVVFSNDRSNGTKETRLDLLKEMIGHEQETMLVKTAKNAYQGAGESQEEEGLDGYTAVGP 140
             | | :|   |   ||    |   | :
Db     272 LVTFGHD-GRGHALTRRRRAKRSPKHHSQR---------------------------- 300

Qy     141 LLARRKRSTGASSHCQKTSLRVNFEDIGWDSWIIAPKEYDAYECKGGCFFPLADDVTPTK 200
             ||:|     : :|:: || |:| |:||: ||:||  | |: | | | |||||  :  |
Db     301 --ARKK-----NKNCRRHSLYVDFSDVGWNDWIVAPPGYQAFYCHGDCPFPLADHLNSTN 353

Qy     201 HAIVQTLVHLKFPTKVGKACCVPTKLSPISILYKDDMGVPTLKYHYEGMSVAECGCR 257
           ||||||||:    : : |||||||:|| ||:|| |:     || :|: | |  ||||
Db     354 HAIVQTLVN-SVNSSIPKACCVPTELSAISMLYLDEYDKVVLK-NYQEMVVEGCGCR 408


RESULT 7
```

```
BMP4_RAT
ID   BMP4_RAT       STANDARD;      PRT;   408 AA.
AC   Q06826;
DT   01-JUN-1994 (Rel. 29, Created)
DT   01-JUN-1994 (Rel. 29, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Bone morphogenetic protein 4 precursor (BMP-4) (BMP-2B).
GN   BMP4 OR BMP-4 OR DVR-4.
OS   Rattus norvegicus (Rat).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX   NCBI_TaxID=10116;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=93385158; PubMed=8373807;
RA   Chen D., Feng J.Q., Feng M., Harris M.A., Mundy G.R., Harris S.E.;
RT   "Cloning and sequence of bone morphogenetic protein 4 cDNA from fetal
RT   rat calvarial cell.";
RL   Biochim. Biophys. Acta 1174:289-292(1993).
CC   -!- FUNCTION: INDUCES CARTILAGE AND BONE FORMATION.
CC   -!- SUBUNIT: HOMODIMER; DISULFIDE-LINKED (BY SIMILARITY).
CC   -!- SUBCELLULAR LOCATION: SECRETED INTO THE EXTRACELLULAR MATRIX.
CC   -!- SIMILARITY: BELONGS TO THE TGF-BETA FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; Z22607; CAA80329.1; -.
DR   PIR; S33173; S33173.
DR   PIR; S38343; S38343.
DR   HSSP; P12643; 3BMP.
DR   InterPro; IPR001839; TGFb.
DR   InterPro; IPR001111; TGFb_N.
DR   Pfam; PF00019; TGF-beta; 1.
DR   Pfam; PF00688; TGFb_propeptide; 1.
DR   ProDom; PD000357; TGFb; 1.
DR   SMART; SM00204; TGFB; 1.
DR   PROSITE; PS00250; TGF_BETA_1; 1.
KW   Signal; Growth factor; Cytokine; Bone; Cartilage; Glycoprotein.
FT   SIGNAL       1     19       POTENTIAL.
FT   PROPEP      20    292
FT   CHAIN      293    408       BONE MORPHOGENETIC PROTEIN 4.
FT   DISULFID    308    373       BY SIMILARITY.
FT   DISULFID    337    405       BY SIMILARITY.
FT   DISULFID    341    407       BY SIMILARITY.
FT   DISULFID    372    372       INTERCHAIN (BY SIMILARITY).
FT   CARBOHYD    144    144       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    209    209       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    350    350       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    365    365       N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ   SEQUENCE   408 AA;  46540 MW;  61E92B4B8D5624F3 CRC64;

  Query Match             25.0%;  Score 342.5;  DB 1;  Length 408;
  Best Local Similarity   35.4%;  Pred. No. 4.5e-23;
  Matches   84;  Conservative   31;  Mismatches   73;  Indels   49;  Gaps    7;

Qy      29 WETLEVSSAVKRWVRADSTTNKNKLEVT----VQSHRESCDTLDISVPPGSKN----LPF 80
           ||| :|| || || |         :|||      ::|:      :  |:| || |     |
Db     213 WETFDVSPAVLRWTREKQPNYGLAIEVTHLHQTRTHQGQHVRISRSLPQGSGNWAQLRPL 272

Qy      81 FVVFSNDRSNGTKETRLDLLKEMIGHEQETMLVKTAKNAYQGAGESQEEEGLDGYTAVGP 140
             | | :|                ||   |:  : || : :
Db     273 LVTFGHDGR--------------GH---TLTRRRAKRSPK------------------ 295

Qy     141 LLARRKRSTGASSHCQKTSLRVNFEDIGWDSWIIAPKEYDAYECKGGCFFPLADDVTPTK 200
                :||   : :|:: || |:| |:||: ||:||  | |: | | | |||||  :  |
Db     296 --HHPQRSRKKNKNCRRHSLYVDFSDVGWNDWIVAPPGYQAFYCHGDCPFPLADHLNSTN 353
```

```
Qy     201 HAIVQTLVHLKFPTKVGKACCVPTKLSPISILYKDDMGVPTLKYHYEGMSVAECGCR 257
           ||||||||:    : : |||||||:|| ||:|| |:     || :|: | |  ||||
Db     354 HAIVQTLVN-SVNSSIPKACCVPTELSAISMLYLDEYDKVVLK-NYQEMVVEGCGCR 408


RESULT 8
BMP4_DAMDA
ID   BMP4_DAMDA     STANDARD;      PRT;   408 AA.
AC   Q29607;
DT   01-NOV-1997 (Rel. 35, Created)
DT   01-NOV-1997 (Rel. 35, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Bone morphogenetic protein 4 precursor (BMP-4).
GN   BMP4.
OS   Dama dama (Fallow deer) (Cervus dama).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Cervoidea;
OC   Cervidae; Cervinae; Cervus.
OX   NCBI_TaxID=30532;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Antler;
RX   MEDLINE=95367593; PubMed=7640308;
RA   Feng J.Q., Chen D., Esparza J., Harris M.A., Mundy G.R.,
RA   Harris S.E.;
RT   "Deer antler tissue contains two types of bone morphogenetic protein
RT   4 mRNA transcripts.";
RL   Biochim. Biophys. Acta 1263:163-168(1995).
CC   -!- FUNCTION: INDUCES CARTILAGE AND BONE FORMATION.
CC   -!- SUBUNIT: HOMODIMER; DISULFIDE-LINKED (BY SIMILARITY).
CC   -!- SIMILARITY: BELONGS TO THE TGF-BETA FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; S79174; AAA80514.1; -.
DR   HSSP; P12643; 3BMP.
DR   InterPro; IPR001839; TGFb.
DR   InterPro; IPR001111; TGFb_N.
DR   Pfam; PF00019; TGF-beta; 1.
DR   Pfam; PF00688; TGFb_propeptide; 1.
DR   ProDom; PD000357; TGFb; 1.
DR   SMART; SM00204; TGFB; 1.
DR   PROSITE; PS00250; TGF_BETA_1; 1.
KW   Signal; Growth factor; Cytokine; Bone; Cartilage; Glycoprotein.
FT   SIGNAL        1     19       POTENTIAL.
FT   PROPEP       20    292       BY SIMILARITY.
FT   CHAIN       293    408       BONE MORPHOGENETIC PROTEIN 4.
FT   DISULFID    308    373       BY SIMILARITY.
FT   DISULFID    337    405       BY SIMILARITY.
FT   DISULFID    341    407       BY SIMILARITY.
FT   DISULFID    372    372       INTERCHAIN (BY SIMILARITY).
FT   CARBOHYD    144    144       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    208    208       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    350    350       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    365    365       N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ   SEQUENCE   408 AA;  46563 MW;  17BA333BB11226FA CRC64;

  Query Match             24.9%;  Score 341;  DB 1;  Length 408;
  Best Local Similarity   35.9%;  Pred. No. 6.1e-23;
  Matches   85;  Conservative   31;  Mismatches   73;  Indels   48;  Gaps    8;

Qy      29 WETLEVSSAVKRWVRADSTTNKNKLEVT----VQSHRESCDTLDISVPPGSKN----LPF 80
           ||| :|| || || |         :|||     ::|:     :  |:| || :     |
Db     212 WETFDVSPAVLRWTREKQPNYGLAIEVTHLHQTRTHQGQHVRISRSLPQGSGDWAQLRPL 271

Qy      81 FVVFSNDRSNGTKETRLDLLKEMIGHEQETMLVKTAKNAYQGAGESQEEEGLDGYTAVGP 140
           | | :|                 ||        |   :: |  | :           |

Db     272 LVTFGHDGR---------------GH---------ALTRHRRAKRSPKHH---------P 298

Qy     141 LLARRKRSTGASSHCQKTSLRVNFEDIGWDSWIIAPKEYDAYECKGGCFFPLADDVTPTK 200
           ||:|     : :|:: || |:| |:||: ||:||  | |: | | | |||||  :  |
Db     299 QRARKK-----NKNCRRHSLYVDFSDVGWNDWIVAPPGYQAFYCHGDCPFPLADHLNSTN 353

Qy     201 HAIVQTLVHLKFPTKVGKACCVPTKLSPISILYKDDMGVPTLKYHYEGMSVAECGCR 257
           ||||||||:    : : |||||||:|| ||:|| |:     || :|: | |  ||||
Db     354 HAIVQTLVN-SVNSSIPKACCVPTELSAISMLYLDEYDKVVLK-NYQEMVVEGCGCR 408


RESULT 9
BMP4_MOUSE
ID   BMP4_MOUSE     STANDARD;      PRT;   408 AA.
AC   P21275;
DT   01-MAY-1991 (Rel. 18, Created)
DT   01-FEB-1994 (Rel. 28, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Bone morphogenetic protein 4 precursor (BMP-4) (BMP-2B).
GN   BMP4 OR BMP-4 OR DVR-4.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Dickinson M.E., van der Meer-De Jong R., Hogan B.L.M.;
RL   Submitted (XXX-1992) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=93282803; PubMed=8507180;
RA   Kurihara T., Kitamura K., Takaoka K., Nakazato H.;
RT   "Murine bone morphogenetic protein-4 gene: existence of multiple
RT   promoters and exons for the 5'-untranslated region.";
RL   Biochem. Biophys. Res. Commun. 192:1049-1056(1993).
RN   [3]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=93365172; PubMed=8358941;
RA   Takaoka K., Yoshikawa H., Hasimoto J., Masuhara K., Miyamoto S.,
RA   Suzuki S., Ono K., Matsui M., Oikawa S., Tsuruoka N.;
RT   "Gene cloning and expression of a bone morphogenetic protein derived
RT   from a murine osteosarcoma.";
RL   Clin. Orthop. Relat. Res. 294:344-352(1993).
RN   [4]
RP   SEQUENCE FROM N.A.
RC   STRAIN=129/Sv; TISSUE=Liver;
RX   MEDLINE=96081880; PubMed=7499338;
RA   Feng J.Q., Chen D., Cooney A.J., Tsai M., Harris M.A., Tsai S.Y.,
RA   Feng M., Mundy G.R., Harris S.E.;
RT   "The mouse bone morphogenetic protein-4 gene. Analysis of promoter
RT   utilization in fetal rat calvarial osteoblasts and regulation by
RT   COUP-TFI orphan receptor.";
RL   J. Biol. Chem. 270:28364-28373(1995).
RN   [5]
RP   SEQUENCE OF 241-408 FROM N.A.
RX   MEDLINE=90228966; PubMed=1970330;
RA   Dickinson M.E., Kobrin M.S., Silan C.M., Kingsley D.M., Justice M.J.,
RA   Miller D.A., Ceci J.D., Lock L.F., Lee A., Buchberg A.M.,
RA   Siracusa L.D., Lyons K.M., Derynck R., Hogan B.L.M., Copeland N.G.,
RA   Jenkins N.A.;
RT   "Chromosomal localization of seven members of the murine TGF-beta
RT   superfamily suggests close linkage to several morphogenetic mutant
RT   loci.";
RL   Genomics 6:505-520(1990).
CC   -!- FUNCTION: INDUCES CARTILAGE AND BONE FORMATION.
CC   -!- SUBUNIT: HOMODIMER; DISULFIDE-LINKED (BY SIMILARITY).
CC   -!- SUBCELLULAR LOCATION: SECRETED INTO THE EXTRACELLULAR MATRIX.
CC   -!- SIMILARITY: BELONGS TO THE TGF-BETA FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
```

```
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; X56848; CAA40179.1; -.
DR   EMBL; S65032; AAB28021.1; -.
DR   EMBL; L47480; AAC37698.1; ALT_INIT.
DR   EMBL; D14814; BAA03555.1; -.
DR   PIR; B34201; B34201.
DR   PIR; S29523; S29523.
DR   HSSP; P12643; 3BMP.
DR   MGD; MGI:88180; Bmp4.
DR   InterPro; IPR001839; TGFb.
DR   InterPro; IPR001111; TGFb_N.
DR   Pfam; PF00019; TGF-beta; 1.
DR   Pfam; PF00688; TGFb_propeptide; 1.
DR   ProDom; PD000357; TGFb; 1.
DR   SMART; SM00204; TGFB; 1.
DR   PROSITE; PS00250; TGF_BETA_1; 1.
KW   Signal; Growth factor; Cytokine; Bone; Cartilage; Glycoprotein.
FT   SIGNAL       1     19       POTENTIAL.
FT   PROPEP      20    292
FT   CHAIN      293    408       BONE MORPHOGENETIC PROTEIN 4.
FT   DISULFID   308    373       BY SIMILARITY.
FT   DISULFID   337    405       BY SIMILARITY.
FT   DISULFID   341    407       BY SIMILARITY.
FT   DISULFID   372    372       INTERCHAIN (BY SIMILARITY).
FT   CARBOHYD   144    144       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   209    209       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   350    350       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   365    365       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CONFLICT   407    407       C -> S (IN REF. 2).
SQ   SEQUENCE   408 AA;  46496 MW;  35053D844624EF9D CRC64;

  Query Match             24.7%;  Score 337.5;  DB 1;  Length 408;
  Best Local Similarity   35.0%;  Pred. No. 1.2e-22;
  Matches   83;  Conservative   32;  Mismatches   73;  Indels   49;  Gaps    7;

Qy      29 WETLEVSSAVKRWVRADSTTNKNKLEVT----VQSHRESCDTLDISVPPGSKN----LPF 80
           ||| :|| || || |          :|||     ::|:     :  |:| || :     |
Db     213 WETFDVSPAVLRWTREKQPNYGLAIEVTHLHQTRTHQGQHVRISRSLPQGSGDWAQLRPL 272

Qy      81 FVVFSNDRSNGTKETRLDLLKEMIGHEQETMLVKTAKNAYQGAGESQEEEGLDGYTAVGP 140
             | | :|                ||   |:  : || : :
Db     273 LVTFGHDGR---------------GH---TLTRRRAKRSPK------------------- 295

Qy     141 LLARRKRSTGASSHCQKTSLRVNFEDIGWDSWIIAPKEYDAYECKGGCFFPLADDVTPTK 200
                 :||   : :|:: || |:| |:||: ||:||  | |: | | | ||||| :  |
Db     296 --HHPQRSRKKNKNCRRHSLYVDFSDVGWNDWIVAPPGYQAFYCHGDCPFPLADHLNSTN 353

Qy     201 HAIVQTLVHLKFPTKVGKACCVPTKLSPISILYKDDMGVPTLKYHYEGMSVAECGCR 257
           ||||||||:    : : |||||||:|| ||:|| |:     || :|: | |  ||||
Db     354 HAIVQTLVN-SVNSSIPKACCVPTELSAISMLYLDEYDKVVLK-NYQEMVVEGCGCR 408


RESULT 10
BMP4_RABIT
ID   BMP4_RABIT     STANDARD;      PRT;   409 AA.
AC   O46576;
DT   15-JUL-1999 (Rel. 38, Created)
DT   15-JUL-1999 (Rel. 38, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Bone morphogenetic protein 4 precursor (BMP-4).
GN   BMP4 OR BMP-4.
OS   Oryctolagus cuniculus (Rabbit).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
OX   NCBI_TaxID=9986;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=New Zealand white; TISSUE=Ocular ciliary epithelium;
RA   Wan X.L., Sears J., Chen S., Sears M.;
RT   "Cloning and expression of BMP-2/-4 from rabbit ocular ciliary
RT   epithelium.";
RL   Submitted (JAN-1998) to the EMBL/GenBank/DDBJ databases.
CC   -!- FUNCTION: INDUCES CARTILAGE AND BONE FORMATION (BY SIMILARITY).
CC   -!- SUBUNIT: HOMODIMER; DISULFIDE-LINKED (BY SIMILARITY).
CC   -!- SUBCELLULAR LOCATION: SECRETED INTO THE EXTRACELLULAR MATRIX.
CC   -!- SIMILARITY: BELONGS TO THE TGF-BETA FAMILY.
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its content  is  in  no  way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; AF042497; AAB97467.1; -.
DR   HSSP; P12643; 3BMP.
DR   InterPro; IPR001839; TGFb.
DR   InterPro; IPR001111; TGFb_N.
DR   Pfam; PF00019; TGF-beta; 1.
DR   Pfam; PF00688; TGFb_propeptide; 1.
DR   ProDom; PD000357; TGFb; 1.
DR   SMART; SM00204; TGFB; 1.
DR   PROSITE; PS00250; TGF_BETA_1; 1.
KW   Signal; Growth factor; Cytokine; Bone; Cartilage; Glycoprotein.
FT   SIGNAL       1     19       POTENTIAL.
FT   PROPEP      20    293       BY SIMILARITY.
FT   CHAIN      294    409       BONE MORPHOGENETIC PROTEIN 4.
FT   DISULFID   309    374       BY SIMILARITY.
FT   DISULFID   338    406       BY SIMILARITY.
FT   DISULFID   342    408       BY SIMILARITY.
FT   DISULFID   373    373       INTERCHAIN (BY SIMILARITY).
FT   CARBOHYD   144    144       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   209    209       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   351    351       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   366    366       N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ   SEQUENCE   409 AA;  46641 MW;  35557561447AD625 CRC64;

  Query Match             24.6%;  Score 336;  DB 1;  Length 409;
  Best Local Similarity   35.4%;  Pred. No. 1.7e-22;
  Matches   84;  Conservative   29;  Mismatches   76;  Indels   48;  Gaps    7;

Qy      29 WETLEVSSAVKRWVRADSTTNKNKLEVT----VQSHRESCDTLDISVPPGSKN----LPF 80
           ||| :|| || || |      :   :|||     ::|:     |  |:  || :     |
Db     213 WETFDVSPAVLRWTREKQPNHGLAVEVTHFHHTRTHQGQHVRLSRSLLQGSGDWAQFRPL 272

Qy      81 FVVFSNDRSNGTKETRLDLLKEMIGHEQETMLVKTAKNAYQGAGESQEEEGLDGYTAVGP 140
             | | :|    |   ||    |  : |                                  |
Db     273 LVTFGHD-GRGHALTRRRRAKRSLKHH------------------------------P 299

Qy     141 LLARRKRSTGASSHCQKTSLRVNFEDIGWDSWIIAPKEYDAYECKGGCFFPLADDVTPTK 200
              ||:|     : :|:: :| |:| |:||: ||:||  | |: | | | |||||     |
Db     300 QRARKK-----NKNCRRHALYVDFSDVGWNDWIVAPPGYQAFYCHGDCPFPLADHFNSTN 354

Qy     201 HAIVQTLVHLKFPTKVGKACCVPTKLSPISILYKDDMGVPTLKYHYEGMSVAECGCR 257
           ||||||||:    : : |||||||:|| ||:|| |:     || :|: | |  ||||
Db     355 HAIVQTLVN-SVNSSIPKACCVPTELSAISMLYLDEYDKVVLK-NYQEMVVEGCGCR 409


RESULT 11
GDF6_BOVIN
ID   GDF6_BOVIN     STANDARD;      PRT;   436 AA.
AC   P55106;
DT   01-OCT-1996 (Rel. 34, Created)
DT   01-OCT-1996 (Rel. 34, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Growth/differentiation factor 6 precursor (GDF-6) (Cartilage-derived
DE   morphogenetic protein 2) (CDMP-2) (Fragment).
GN   GDF6 OR CDMP2.
OS   Bos taurus (Bovine).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;
OC   Bovidae; Bovinae; Bos.
```

```
OX   NCBI_TaxID=9913;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Articular cartilage;
RX   MEDLINE=95050604; PubMed=7961761;
RA   Chang S., Hoang B., Thomas J.T., Vukicevic S., Luyten F.P.,
RA   Ryba N.J.P., Kozak C.A., Reddi A.H., Moos M.;
RT   "Cartilage-derived morphogenetic proteins. New members of the
RT   transforming growth factor-beta superfamily predominantly expressed
RT   in long bones during human embryonic development.";
RL   J. Biol. Chem. 269:28227-28234(1994).
CC   -!- SUBUNIT: HOMODIMER; DISULFIDE-LINKED (BY SIMILARITY).
CC   -!- SIMILARITY: BELONGS TO THE TGF-BETA FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; U13661; AAA61416.1; -.
DR   HSSP; P18075; 1BMP.
DR   InterPro; IPR001839; TGFb.
DR   InterPro; IPR001111; TGFb_N.
DR   Pfam; PF00019; TGF-beta; 1.
DR   Pfam; PF00688; TGFb_propeptide; 1.
DR   ProDom; PD000357; TGFb; 1.
DR   SMART; SM00204; TGFB; 1.
DR   PROSITE; PS00250; TGF_BETA_1; 1.
KW   Growth factor; Cytokine; Glycoprotein.
FT   NON_TER       1      1
FT   PROPEP       <1    316       POTENTIAL.
FT   CHAIN       317    436       GROWTH/DIFFERENTIATION FACTOR 6.
FT   DISULFID    335    401       BY SIMILARITY.
FT   DISULFID    364    433       BY SIMILARITY.
FT   DISULFID    368    435       BY SIMILARITY.
FT   DISULFID    400    400       INTERCHAIN (BY SIMILARITY).
FT   CARBOHYD     27     27       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD     89     89       N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ   SEQUENCE   436 AA;  47873 MW;  B0688E12EF8AE91D CRC64;

  Query Match             24.1%;  Score 330;  DB 1;  Length 436;
  Best Local Similarity   36.7%;  Pred. No. 6.3e-22;
  Matches   81;  Conservative   23;  Mismatches   67;  Indels   50;  Gaps    7;

Qy      72 PPGSKNLPF------------FVVFSNDRSNGTKETRLDLLKEMIGHEQETMLVKTAKNA 119
           ||  ::| |             ||||  :    ::|    ::| :|   | :
Db     231 PPDLRSLGFGRRVRTPQERALLVVFSRSQ----RKTLFAEMREQLGSATEVV-------- 278

Qy     120 YQGAGESQEEEG---------LDGYTAVG---PLLARRKRSTGASSH-----------CQ 156
             | |   | |           |    |   ||:|:  || |           |
Db     279 --GPGGGAEGSGPPPPPPPPPPSGTPDAGLWSPSPGRRRRTAFASRHGKRHGKKSRLRCS 336

Qy     157 KTSLRVNFEDIGWDSWIIAPKEYDAYECKGGCFFPLADDVTPTKHAIVQTLVHLKFPTKV 216
           |  | |||:::|||  ||||| ||:|| |:| | |||   : || |||:|||::   |
Db     337 KKPLHVNFKELGWDDWIIAPLEYEAYHCEGVCDFPLRSHLEPTNHAIIQTLMNSMDPGST 396

Qy     217 GKACCVPTKLSPISILYKDDMGVPTLKYHYEGMSVAECGCR 257
           :|||||||:||||||  | |   :   || | |  ||||
Db     397 PPSCCVPTKLTPISILY-IDAGNNVVYNEYEEMVVESCGCR 436


RESULT 12
BMP4_CHICK
ID   BMP4_CHICK     STANDARD;      PRT;   405 AA.
AC   Q90752;
DT   01-NOV-1997 (Rel. 35, Created)
DT   01-NOV-1997 (Rel. 35, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Bone morphogenetic protein 4 precursor (BMP-4).
GN   BMP4 OR BMP-4.
OS   Gallus gallus (Chicken).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC   Gallus.
OX   NCBI_TaxID=9031;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=White leghorn;
RX   MEDLINE=94163974; PubMed=8119128;
RA   Francis P.H., Richardson M.K., Brickell P.M., Tickle C.;
RT   "Bone morphogenetic proteins and a signalling pathway that controls
RT   patterning in the developing chick limb.";
RL   Development 120:209-218(1994).
RN   [2]
RP   FUNCTION.
RX   MEDLINE=99128179; PubMed=9927590;
RA   Pizette S., Niswander L.;
RT   "BMPs negatively regulate structure and function of the limb apical
RT   ectodermal ridge.";
RL   Development 126:883-894(1999).
CC   -!- FUNCTION: NEGATIVELY REGULATES THE STRUCTURE AND FUNCTION OF THE
CC       LIMB APICAL ECTODERMAL RIDGE.
CC   -!- SUBUNIT: HOMODIMER; DISULFIDE-LINKED (BY SIMILARITY).
CC   -!- SIMILARITY: BELONGS TO THE TGF-BETA FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; X75915; CAA53514.1; -.
DR   HSSP; P12643; 3BMP.
DR   InterPro; IPR001839; TGFb.
DR   InterPro; IPR001111; TGFb_N.
DR   Pfam; PF00019; TGF-beta; 1.
DR   Pfam; PF00688; TGFb_propeptide; 1.
DR   ProDom; PD000357; TGFb; 1.
DR   SMART; SM00204; TGFB; 1.
DR   PROSITE; PS00250; TGF_BETA_1; 1.
KW   Signal; Growth factor; Cytokine; Bone; Cartilage; Glycoprotein.
FT   SIGNAL        1     19       POTENTIAL.
FT   PROPEP       20    291       BY SIMILARITY.
FT   CHAIN       292    405       BONE MORPHOGENETIC PROTEIN 4.
FT   DISULFID    305    370       BY SIMILARITY.
FT   DISULFID    334    402       BY SIMILARITY.
FT   DISULFID    338    404       BY SIMILARITY.
FT   DISULFID    369    369       INTERCHAIN (BY SIMILARITY).
FT   CARBOHYD    144    144       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    208    208       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    347    347       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    362    362       N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ   SEQUENCE   405 AA;  46057 MW;  544302DBA0A40F81 CRC64;

  Query Match             23.9%;  Score 327.5;  DB 1;  Length 405;
  Best Local Similarity   33.2%;  Pred. No. 9.6e-22;
  Matches   79;  Conservative   36;  Mismatches   70;  Indels   53;  Gaps    7;

Qy      29 WETLEVSSAVKRWVRADSTTNKNKLEVT----VQSHRESCDTLDISVPPG-----SKNLP 79
           ||| :|| || || :     :   :|||     |:|:     :  |:| |     ::  |
Db     212 WETFDVSPAVIRWTKDKQPNHGLVIEVTHLHQAQTHQGKHVRISRSLPQGHGGDWAQLRP 271

Qy      80 FFVVFSNDRSNGTKETRLDLLKEMIGHEQETMLVKTAKNAYQGAGESQEEEGLDGYTAVG 139
             | | :|                 ||     | : |: : :  |  : ::
Db     272 LLVTFGHDGR--------------GH----ALTRRARRSPKHHGSRKNKK--------- 303

Qy     140 PLLARRKRSTGASSHCQKTSLRVNFEDIGWDSWIIAPKEYDAYECKGGCFFPLADDVTPT 199
                       :|:: :| |:| |:||: ||:||  | |: | | | |||||  :  |
Db     304 --------------NCRRHALYVDFSDVGWNDWIVAPPGYQAFYCHGDCPFPLADHLNST 349

Qy     200 KHAIVQTLVHLKFPTKVGKACCVPTKLSPISILYKDDMGVPTLKYHYEGMSVAECGCR 257
```

```
          ||||||||:    : : |||||||:|| ||:|| |:     || :|: | |  ||||
Db    350 NHAIVQTLVN-SVNSSIPKACCVPTELSAISMLYLDEYDKVVLK-NYQEMVVEGCGCR 405


RESULT 13
DECA_DROME
ID   DECA_DROME     STANDARD;      PRT;   588 AA.
AC   P07713; P91651;
DT   01-APR-1988 (Rel. 07, Created)
DT   01-APR-1988 (Rel. 07, Last sequence update)
DT   15-DEC-1998 (Rel. 37, Last annotation update)
DE   Decapentaplegic protein precursor (DPP-C protein).
GN   DPP.
OS   Drosophila melanogaster (Fruit fly).
OC   Eukaryota; Metazoa; Arthropoda; Mandibulata; Pancrustacea; Hexapoda;
OC   Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
OC   Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
OX   NCBI_TaxID=7227;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=87090408; PubMed=3467201;
RA   Padgett R.W., St Johnston R.D., Gelbart W.M.;
RT   "A transcript from a Drosophila pattern gene predicts a protein
RT   homologous to the transforming growth factor-beta family.";
RL   Nature 325:81-84(1987).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=DP CN BW;
RX   MEDLINE=97225213; PubMed=9071586;
RA   Richter B., Long M., Lewontin R.C., Nitasaka E.;
RT   "Nucleotide variation and conservation at the dpp locus, a gene
RT   controlling early development in Drosophila.";
RL   Genetics 145:311-323(1997).
RN   [3]
RP   CHARACTERIZATION, AND SEQUENCE OF 457-476.
RX   MEDLINE=90258853; PubMed=1692958;
RA   Panganiban G.E.F., Rashka K.E., Neitzel M.D., Hoffmann F.M.;
RT   "Biochemical characterization of the Drosophila dpp protein, a member
RT   of the transforming growth factor beta family of growth factors.";
RL   Mol. Cell. Biol. 10:2669-2677(1990).
CC   -!- FUNCTION: ACTS AS AN EXTRACELLULAR MORPHOGEN TO ESTABLISH AT LEAST
CC       TWO CELLULAR RESPONSE THRESHOLDS WITHIN THE DORSAL HALF OF THE
CC       DROSOPHILA EMBRYO. REQUIRED FOR THE PROPER DEVELOPMENT OF THE
CC       EMBRYONIC DORSAL HYPODERM, FOR VIABILITY OF LARVAE AND FOR CELL
CC       VIABILITY OF THE EPITHELIAL CELLS IN THE IMAGINAL DISKS. ACTS
CC       TOGETHER WITH SCW.
CC   -!- SUBUNIT: HETERODIMERS OF SCW/DPP ARE THE ACTIVE SUBUNIT, DPP/DPP
CC       HOMODIMERS ELICIT A BASAL RESPONSE AND SCW/SCW HOMODIMERS ALONE
CC       ARE INEFFECTIVE IN SPECIFYING A DORSAL PATTERN.
CC   -!- TISSUE SPECIFICITY: EXPRESSED IN THE IMAGINAL DISCS ASSOCIATED
CC       WITH ESTABLISHMENT OF THE PROXIMAL-DISTAL AXIS OF THE APPENDAGES,
CC       AND MIDGUT MESODERM.
CC   -!- SIMILARITY: BELONGS TO THE TGF-BETA FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; M30116; AAA28482.1; -.
DR   EMBL; U63857; AAC47552.1; -.
DR   PIR; A26158; A26158.
DR   HSSP; P12643; 3BMP.
DR   FlyBase; FBgn0000490; dpp.
DR   InterPro; IPR001839; TGFb.
DR   InterPro; IPR001111; TGFb_N.
DR   Pfam; PF00019; TGF-beta; 1.
DR   Pfam; PF00688; TGFb_propeptide; 1.
DR   ProDom; PD000357; TGFb; 1.
DR   SMART; SM00204; TGFB; 1.
DR   PROSITE; PS00250; TGF_BETA_1; 1.
KW   Growth factor; Developmental protein; Differentiation; Glycoprotein;
KW   Signal.
FT   SIGNAL        1     15       POTENTIAL.
FT   PROPEP       16    456
FT   CHAIN       457    588       DECAPENTAPLEGIC PROTEIN.
FT   DISULFID    487    553       BY SIMILARITY.
FT   DISULFID    516    585       BY SIMILARITY.
FT   DISULFID    520    587       BY SIMILARITY.
FT   DISULFID    552    552       INTERCHAIN (BY SIMILARITY).
FT   CARBOHYD    120    120       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    342    342       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    377    377       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    529    529       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   VARIANT      59     59       V -> G (IN STRAIN DP CN BW).
FT   VARIANT     121    121       K -> M (IN STRAIN DP CN BW).
FT   VARIANT     473    474       HA -> QP (IN STRAIN DP CN BW).
SQ   SEQUENCE   588 AA;  65850 MW;  3D986A7DF5DF666B CRC64;

  Query Match             23.8%;  Score 325;  DB 1;  Length 588;
  Best Local Similarity   30.5%;  Pred. No. 2.6e-21;
  Matches   92;  Conservative   45;  Mismatches  115;  Indels   50;  Gaps    9;

Qy       4 EDSETWDQATGTKTFLVSQDIR----DEGWETLEV----SSAVKRWVRADSTTNKNKLEV 55
           :||:  |:       | :  |::    ||  :  |:     :  :: | : |: |: : :|
Db     289 KDSKIDDRFPHHHRFRLHFDVKSIPADEKLKAAELQLTRDALSQQVVASRSSANRTRYQV 348

Qy      56 T--------VQSHRE-------------SCDTLDISVPPG--------SKNLPFFVVFS 85
             |:  ||                  | ||: : | |           :|    |
Db     349 LVYDITRVGVRGQREPSYLLLDTKTVRLNSTDTVSLDVQPAVDRWLASPQRNYGLLVEVR 408

Qy      86 NDRS-------NGTKETRLDLLKEMIGHEQETMLVKTAKNAYQGAGESQEEEGLDGYTAV 138
             ||       :       |  |   |:|  :   |    :: |   ::  | :|
Db     409 TVRSLKPAPHHHVRLRRSADEAHERWQHKQPLLFTYTDDGRHK-ARSIRDVSGGEGGGKG 467

Qy     139 GPLLARRKRSTGASSH---CQKTSLRVNFEDIGWDSWIIAPKEYDAYECKGGCFFPLADD 195
           |      :| |   :|   |:: || |:| |:||| ||:||  |||| | | | |||||
Db     468 GRNKRHARRPTRRKNHDDTCRRHSLYVDFSDVGWDDWIVAPLGYDAYYCHGKCPFPLADH 527

Qy     196 VTPTKHAIVQTLVHLKFPTKVGKACCVPTKLSPISILYKDDMGVPTLKYHYEGMSVAECG 255
             | ||:||||||:   | || |||||||:|  :::|| :|     || :|: |:|  ||
Db     528 FNSTNHAVVQTLVNNMNPGKVPKACCVPTQLDSVAMLYLNDQSTVVLK-NYQEMTVVGCG 586

Qy     256 CR 257
           ||
Db     587 CR 588


RESULT 14
DECA_DROPS
ID   DECA_DROPS     STANDARD;      PRT;   621 AA.
AC   P91699;
DT   15-DEC-1998 (Rel. 37, Created)
DT   15-DEC-1998 (Rel. 37, Last sequence update)
DT   15-DEC-1998 (Rel. 37, Last annotation update)
DE   Decapentaplegic protein precursor (DPP-C protein).
GN   DPP.
OS   Drosophila pseudoobscura (Fruit fly).
OC   Eukaryota; Metazoa; Arthropoda; Mandibulata; Pancrustacea; Hexapoda;
OC   Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
OC   Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
OX   NCBI_TaxID=7237;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=97225212; PubMed=9071585;
RA   Newfeld S.J., Padgett R.W., Findley S.D., Richter B.G., Sanicola M.,
RA   de Cuevas M., Gelbart W.M.;
RT   "Molecular evolution at the decapentaplegic locus in Drosophila.";
RL   Genetics 145:297-309(1997).
CC   -!- FUNCTION: ACTS AS AN EXTRACELLULAR MORPHOGEN TO ESTABLISH AT LEAST
CC       TWO CELLULAR RESPONSE THRESHOLDS WITHIN THE DORSAL HALF OF THE
CC       DROSOPHILA EMBRYO. REQUIRED FOR THE PROPER DEVELOPMENT OF THE
CC       EMBRYONIC DORSAL HYPODERM, FOR VIABILITY OF LARVAE AND FOR CELL
```

```
CC       VIABILITY OF THE EPITHELIAL CELLS IN THE IMAGINAL DISKS. ACTS
CC       TOGETHER WITH SCW (BY SIMILARITY).
CC   -!- SUBUNIT: HETERODIMERS OF SCW/DPP ARE THE ACTIVE SUBUNIT, DPP/DPP
CC       HOMODIMERS ELICIT A BASAL RESPONSE AND SCW/SCW HOMODIMERS ALONE
CC       ARE INEFFECTIVE IN SPECIFYING A DORSAL PATTERN (BY SIMILARITY).
CC   -!- TISSUE SPECIFICITY: EXPRESSED IN THE IMAGINAL DISCS ASSOCIATED
CC       WITH ESTABLISHMENT OF THE PROXIMAL-DISTAL AXIS OF THE APPENDAGES,
CC       AND MIDGUT MESODERM (BY SIMILARITY).
CC   -!- SIMILARITY: BELONGS TO THE TGF-BETA FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; U63856; AAC47553.1; -.
DR   HSSP; P12643; 3BMP.
DR   FlyBase; FBgn0018547; Dpse\dpp.
DR   InterPro; IPR001839; TGFb.
DR   InterPro; IPR001111; TGFb_N.
DR   Pfam; PF00019; TGF-beta; 1.
DR   Pfam; PF00688; TGFb_propeptide; 1.
DR   ProDom; PD000357; TGFb; 1.
DR   SMART; SM00204; TGFB; 1.
DR   PROSITE; PS00250; TGF_BETA_1; 1.
KW   Growth factor; Developmental protein; Differentiation; Glycoprotein;
KW   Signal.
FT   SIGNAL       1     15       POTENTIAL.
FT   PROPEP      16    479       BY SIMILARITY.
FT   CHAIN      480    621       DECAPENTAPLEGIC PROTEIN.
FT   DISULFID   520    586       BY SIMILARITY.
FT   DISULFID   549    618       BY SIMILARITY.
FT   DISULFID   553    620       BY SIMILARITY.
FT   DISULFID   585    585       INTERCHAIN (BY SIMILARITY).
FT   DOMAIN      45     57       POLY-ALA.
FT   DOMAIN      88     92       POLY-ASN.
FT   DOMAIN      95    104       POLY-ASN.
FT   DOMAIN     163    173       POLY-GLN.
FT   DOMAIN     483    495       POLY-GLY.
FT   CARBOHYD   149    149       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   365    365       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   400    400       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   562    562       N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ   SEQUENCE   621 AA;  69038 MW;  3FD7141FB5509651 CRC64;

  Query Match             23.7%;  Score 324.5;  DB 1;  Length 621;
  Best Local Similarity   34.4%;  Pred. No. 3e-21;
  Matches   86;  Conservative   31;  Mismatches   76;  Indels   57;  Gaps    9;

Qy      31 TLEVSSAVKRWVRADSTTNKN-KLEVTVQSHRESCDTLDISVPPGSKNLPFFVVFSNDRS 89
           :|:|  || ||:   :|  ||  | | |:: |        |: |   :
Db     406 SLDVQPAVDRWL---ATPQKNYGLLVEVRTMR--------SLKPAPHH------------ 442

Qy      90 NGTKETRL----DLLKEMIGHEQETMLVKTAKNAYQ--------------GAGESQEEE 130
                ||    |   |  |:|  :   |    ::                ||||  :
Db     443 ----HVRLRRSADEAHEQWQHKQPLLFAYTDDGRHKARSIRDVSGGGGGGGGAGEGGKGN 498

Qy     131 GLDGYTAVGPLLARRKRSTGASSH---CQKTSLRVNFEDIGWDSWIIAPKEYDAYECKGG 187
           |       |     ::|     :|   |:: || |:| |:||| ||:||  |||| | |
Db     499 G------GGRNRRHQRRPARRKNHEETCRRHSLYVDFADVGWDDWIVAPPGYDAYYCHGK 552

Qy     188 CFFPLADDVTPTKHAIVQTLVHLKFPTKVGKACCVPTKLSPISILYKDDMGVPTLKYHYE 247
           | |||||    | ||:|||||:   | || |||||||:|  :::|| :|     || :|:
Db     553 CPFPLADHFNSTNHAVVQTLVNNLNPGKVPKACCVPTQLDSVAMLYLNDQSTVVLK-NYQ 611

Qy     248 GMSVAECGCR 257
           |:|  ||||
Db     612 EMTVVGCGCR 621
```

```
RESULT 15
DECA_DROSI
ID   DECA_DROSI     STANDARD;      PRT;   593 AA.
AC   P91706;
DT   15-DEC-1998 (Rel. 37, Created)
DT   15-DEC-1998 (Rel. 37, Last sequence update)
DT   15-DEC-1998 (Rel. 37, Last annotation update)
DE   Decapentaplegic protein precursor (DPP-C protein).
GN   DPP.
OS   Drosophila simulans (Fruit fly).
OC   Eukaryota; Metazoa; Arthropoda; Mandibulata; Pancrustacea; Hexapoda;
OC   Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
OC   Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
OX   NCBI_TaxID=7240;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=NET DP1;
RX   MEDLINE=97225212; PubMed=9071585;
RA   Newfeld S.J., Padgett R.W., Findley S.D., Richter B.G., Sanicola M.,
RA   de Cuevas M., Gelbart W.M.;
RT   "Molecular evolution at the decapentaplegic locus in Drosophila.";
RL   Genetics 145:297-309(1997).
CC   -!- FUNCTION: ACTS AS AN EXTRACELLULAR MORPHOGEN TO ESTABLISH AT LEAST
CC       TWO CELLULAR RESPONSE THRESHOLDS WITHIN THE DORSAL HALF OF THE
CC       DROSOPHILA EMBRYO. REQUIRED FOR THE PROPER DEVELOPMENT OF THE
CC       EMBRYONIC DORSAL HYPODERM, FOR VIABILITY OF LARVAE AND FOR CELL
CC       VIABILITY OF THE EPITHELIAL CELLS IN THE IMAGINAL DISKS. ACTS
CC       TOGETHER WITH SCW (BY SIMILARITY).
CC   -!- SUBUNIT: HETERODIMERS OF SCW/DPP ARE THE ACTIVE SUBUNIT, DPP/DPP
CC       HOMODIMERS ELICIT A BASAL RESPONSE AND SCW/SCW HOMODIMERS ALONE
CC       ARE INEFFECTIVE IN SPECIFYING A DORSAL PATTERN (BY SIMILARITY).
CC   -!- TISSUE SPECIFICITY: EXPRESSED IN THE IMAGINAL DISCS ASSOCIATED
CC       WITH ESTABLISHMENT OF THE PROXIMAL-DISTAL AXIS OF THE APPENDAGES,
CC       AND MIDGUT MESODERM (BY SIMILARITY).
CC   -!- SIMILARITY: BELONGS TO THE TGF-BETA FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; U63854; AAC47554.1; -.
DR   HSSP; P12643; 3BMP.
DR   FlyBase; FBgn0015673; Dsim\dpp.
DR   InterPro; IPR001839; TGFb.
DR   InterPro; IPR001111; TGFb_N.
DR   Pfam; PF00019; TGF-beta; 1.
DR   Pfam; PF00688; TGFb_propeptide; 1.
DR   ProDom; PD000357; TGFb; 1.
DR   SMART; SM00204; TGFB; 1.
DR   PROSITE; PS00250; TGF_BETA_1; 1.
KW   Growth factor; Developmental protein; Differentiation; Glycoprotein;
KW   Signal.
FT   SIGNAL       1     15       POTENTIAL.
FT   PROPEP      16    461       BY SIMILARITY.
FT   CHAIN      462    593       DECAPENTAPLEGIC PROTEIN.
FT   DISULFID   492    558       BY SIMILARITY.
FT   DISULFID   521    590       BY SIMILARITY.
FT   DISULFID   525    592       BY SIMILARITY.
FT   DISULFID   557    557       INTERCHAIN (BY SIMILARITY).
FT   CARBOHYD   122    122       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   347    347       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   382    382       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   534    534       N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ   SEQUENCE   593 AA;  66248 MW;  F0BDB21209F44380 CRC64;

  Query Match             23.6%;  Score 323;  DB 1;  Length 593;
  Best Local Similarity   30.5%;  Pred. No. 3.9e-21;
  Matches   92;  Conservative   45;  Mismatches  115;  Indels   50;  Gaps    9;
```

```
Qy       4 EDSETWDQATGTKTFLVSQDIR----DEGWETLEV----SSAVKRWVRADSTTNKNKLEV 55
           :||:  |:      | :  |::    ||  :  |:     :   : | : |: |: : :|
Db     294 KDSKIDDRFPHHHRFRLHFDVKSIPADEKLKAAELQLTRDALSPQVVASRSSANRTRYQV 353

Qy      56 T--------VQSHRE-------------SCDTLDISVPPG--------SKNLPFFVVFS 85
                  |:  ||               | ||: : | |          :|    |
Db     354 LVYDITRVGVRGQREPSYLLLDTKTVRLNSTDTVSLDVQPAVDRWLASPQRNYGLLVEVR 413

Qy      86 NDRS-------NGTKETRLDLLKEMIGHEQETMLVKTAKNAYQGAGESQEEEGLDGYTAV 138
             ||       :       |   |   |:|  :   |    :: |   ::  | :|
Db     414 TVRSLKPAPHHHVRLRRSADEAHERWQHKQPLLFTYTDDGRHK-ARSIRDVSGGEGGGKG 472

Qy     139 GPLLARRKRSTGASSH---CQKTSLRVNFEDIGWDSWIIAPKEYDAYECKGGCFFPLADD 195
           |    : :| |   :|   |:: || |:| |:|:|| ||:||  |||| | | | |||||
Db     473 GRNKRQPRRPTRRKNHDDTCRRHSLYVDFSDVGWDDWIVAPLGYDAYYCHGKCPFPLADH 532

Qy     196 VTPTKHAIVQTLVHLKFPTKVGKACCVPTKLSPISILYKDDMGVPTLKYHYEGMSVAECG 255
              | ||:|||||:   | || |||||||:|  :::|| :|     || :|: |:|  ||
Db     533 FNSTNHAVVQTLVNNMNPGKVPKACCVPTQLDSVAMLYLNDQSTVVLK-NYQEMTVVGCG 591

Qy     256 CR 257
           ||
Db     592 CR 593
```

Search completed: April 2, 2003, 14:50:07
Job time : 28 secs

GenCore version 5.1.4_p5_4578

OM protein - protein search, using sw model

Run on: April 2, 2003, 14:47:35 ; Search time 44 Seconds
(without alignments)
561.512 Million cell updates/sec

Title: US-10-002-278-9
Perfect score: 1368
Sequence: 1 DVLEDSETWDQATGTKTFLV.....GVPTLKYHYEGMSVAECGCR 257

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_73:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	808.5	59.1	427	2	A40735	TGF beta homolog d
2	345	25.2	408	1	BMHU4	bone morphogenetic
3	342.5	25.0	408	2	S38343	bone morphogenetic
4	341	24.9	408	2	S58791	bone morphogenetic
5	337.5	24.7	420	2	I49541	bone morphogenetic
6	330	24.1	436	2	B55452	cartilage-derived
7	327.5	23.9	405	2	I50608	bone morphogenetic
8	327.5	23.9	408	2	JH0801	bone morphogenetic
9	326.5	23.9	400	2	A49147	bone morphogenetic
10	325	23.8	588	2	A26158	decapentaplegic pr
11	319	23.3	398	2	JH0688	bone morphogenetic
12	318	23.2	398	2	JH0687	bone morphogenetic
13	316.5	23.1	401	2	JH0689	bone morphogenetic
14	314.5	23.0	430	2	JQ1184	osteogenic protein
15	314	23.0	393	2	S37073	bone morphogenetic
16	313	22.9	394	2	S45355	bone morphogenetic
17	312.5	22.8	454	1	BMHU5	bone morphogenetic
18	312	22.8	125	2	S43295	bone morphogenetic
19	311	22.7	396	1	BMHU2	bone morphogenetic
20	310.5	22.7	313	2	I51284	bone morphogenetic
21	310.5	22.7	461	2	S52408	SPDVR1 protein - s
22	308	22.5	501	2	A55452	cartilage-derived
23	307	22.4	353	2	I50607	bone morphogenetic
24	307	22.4	501	2	JC2347	growth/differentia
25	305.5	22.3	431	1	BMHU7	bone morphogenetic
26	305.5	22.3	452	2	I49542	bone morphogenetic
27	305.5	22.3	495	2	S43294	bone morphogenetic
28	301	22.0	402	2	A45056	osteogenic protein
29	301	22.0	510	2	A54798	Vg-1-related prote
30	298	21.8	513	1	BMHU6	bone morphogenetic
31	297	21.7	426	2	JH0690	bone morphogenetic
32	294	21.5	151	2	S43296	bone morphogenetic
33	290	21.2	360	2	A29619	Vg1 embryonic grow
34	288.5	21.1	365	2	T43286	cet-1 protein - Ca
35	285	20.8	354	2	S29718	gene nodal protein
36	281	20.5	207	2	S37618	vgr protein - rat
37	274.5	20.1	455	2	A43918	TGF-beta-related p
38	266.5	19.5	476	2	JC4646	bone morphogenetic
39	259.5	19.0	366	2	A46607	growth/differentia
40	255	18.6	360	2	I53032	bone morphogenetic
41	252.5	18.5	478	2	JC4838	bone morphogenetic
42	251.5	18.4	366	2	A45402	transforming growt
43	241.5	17.7	472	1	BMHU3	bone morphogenetic
44	229	16.7	357	2	A39364	GDF-1 embryonic gr
45	228	16.7	413	2	JC4862	activin beta-A cha

ALIGNMENTS

RESULT 1
A40735
TGF beta homolog dsl-1 - chicken
C;Species: Gallus gallus (chicken)
C;Date: 21-Sep-1993 #sequence_revision 18-Nov-1994 #text_change 21-Jul-2000
C;Accession: A40735
R;Basler, K.; Edlund, T.; Jessell, T.M.; Yamada, T.
Cell 73, 687-702, 1993
A;Title: Control of cell pattern in the neural tube: regulation of cell differentiati
A;Reference number: A40735; MUID:93272310; PMID:7916656
A;Accession: A40735
A;Status: preliminary
A;Molecule type: nucleic acid
A;Residues: 1-427 <BAS>
A;Cross-references: GB:L12032; NID:g304379; PIDN:AAA48752.1; PID:g304380
A;Experimental source: spinal cord
A;Note: sequence extracted from NCBI backbone (NCBIN:132680, NCBIP:132681)
C;Superfamily: inhibin

Query Match 59.1%; Score 808.5; DB 2; Length 427;
Best Local Similarity 61.9%; Pred. No. 4e-62;
Matches 161; Conservative 31; Mismatches 59; Indels 9; Gaps 5;

```
Qy      1 DVLEDSETWDQATGTKTFLVSQDIRDEGWETLEVSSAVKRWVRADSTTNKNKLEVTVQSH 60
          ||| | : |:    ||: ||| |:| ||| ||||||||||:||    |||||| ::|
Db    174 DVL-DGDHWENKESTKSLLVSHSIQDCGWEMFEVSSAVKRWVKADKMKTKNKLEVVIESK 232

Qy     61 RES---CDTLDISVPPGSKNLPFFVVFSNDRSNGTKETRLDLLKEMIGHEQETMLVKTAK 117
            |   | |||:|   :|||| :|||||||||||||::: |:||| ||||::| |  |
Db    233 DLSGFPCGKLDITVTHDTKNLPLLIVFSNDRSNGTKETKVE-LREMIVHEQESVLNKLGK 291

Qy    118 NAYQGAGESQEEEGLDGYTAVGPLLARRKRSTGASSHCQKTSLRVNFEDIGWDSWIIAPK 177
          |     | :||: :          :| ||| || :||::||| |||::|||||||||||
Db    292 NDSSSEEEQREEKAI---ARPRQHSSRSKRSIGA-NHCRRTSLHVNFKEIGWDSWIIAPK 347

Qy    178 EYDAYECKGGCFFPLADDVTPTKHAIVQTLVHLKFPTKVGKACCVPTKLSPISILYKDDM 237
          :|:|:|||||||||| |:|||||||||||||||: | |  |||||||||  ||||||||
Db    348 DYEAFECKGGCFFPLTDNVTPTKHAIVQTLVHLQNPKKASKACCVPTKLDAISILYKDDA 407

Qy    238 GVPTLKYHYEGMSVAECGCR 257
          ||||| |:|||| |||||||
Db    408 GVPTLIYNYEGMKVAECGCR 427
```

RESULT 2
BMHU4
bone morphogenetic protein 4 precursor - human
N;Alternate names: bone morphogenetic protein 2B
C;Species: Homo sapiens (man)
C;Date: 16-Sep-1992 #sequence_revision 03-Aug-1995 #text_change 18-Jun-1999
C;Accession: C37278
R;Wozney, J.M.; Rosen, V.; Celeste, A.J.; Mitsock, L.M.; Whitters, M.J.; Kriz, R.W.;

```
Science 242, 1528-1534, 1988
A;Title: Novel regulators of bone formation: molecular clones and activities.
A;Reference number: A37278; MUID:89072730; PMID:3201241
A;Accession: C37278
A;Molecule type: mRNA
A;Residues: 1-408 <WO3>
A;Cross-references: GB:M22490; NID:g179503; PIDN:AAA51835.1; PID:g179504
C;Genetics:
A;Gene: GDB:BMP4; BMP2B
A;Cross-references: GDB:125205; OMIM:112262
A;Map position: 14q22-14q23
C;Superfamily: inhibin
C;Keywords: bone; glycoprotein
F;1-19/Domain: signal sequence #status predicted <SIG>
F;20-292/Domain: propeptide #status predicted <PRO>
F;293-408/Product: bone morphogenetic protein 4 #status predicted <MAT>
F;143,208,350,365/Binding site: carbohydrate (Asn) (covalent) #status predicted

  Query Match             25.2%;  Score 345;  DB 1;  Length 408;
  Best Local Similarity   35.9%;  Pred. No. 4.5e-22;
  Matches   85;  Conservative   28;  Mismatches   76;  Indels   48;  Gaps    7;

Qy      29 WETLEVSSAVKRWVRADSTTNKNKLEVT----VQSHRESCDTLDISVPPGSKN----LPF 80
           ||| :|| || || |         :|||     ::|:     :  |:| || |     |
Db     212 WETFDVSPAVLRWTREKQPNYGLAIEVTHLHQTRTHQGQHVRISRSLPQGSGNWAQLRPL 271

Qy      81 FVVFSNDRSNGTKETRLDLLKEMIGHEQETMLVKTAKNAYQGAGESQEEEGLDGYTAVGP 140
            | | :|   |   ||    |    |  :
Db     272 LVTFGHD-GRGHALTRRRRAKRSPKHHSQR------------------------------ 300

Qy     141 LLARRKRSTGASSHCQKTSLRVNFEDIGWDSWIIAPKEYDAYECKGGCFFPLADDVTPTK 200
             ||:|     : :|:: || |:| |:||: ||:||  | |: | | | ||||| :  |
Db     301 --ARKK-----NKNCRRHSLYVDFSDVGWNDWIVAPPGYQAFYCHGDCPFPLADHLNSTN 353

Qy     201 HAIVQTLVHLKFPTKVGKACCVPTKLSPISILYKDDMGVPTLKYHYEGMSVAECGCR 257
           ||||||||:    : : |||||||:|| ||:|| |:     || :|: | |  ||||
Db     354 HAIVQTLVN-SVNSSIPKACCVPTELSAISMLYLDEYDKVVLK-NYQEMVVEGCGCR 408


RESULT 3
S38343
bone morphogenetic protein 4 - rat
C;Species: Rattus norvegicus (Norway rat)
C;Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 16-Jul-1999
C;Accession: S38343; S33173
R;Chen, D.; Feng, J.Q.; Feng, M.; Harris, M.A.; Mundy, G.R.; Harris, S.E.
Biochim. Biophys. Acta 1174, 289-292, 1993
A;Title: Cloning and sequence of bone morphogenetic protein 4 cDNA from fetal rat calvar
A;Reference number: S38343; MUID:93385158; PMID:8373807
A;Accession: S38343
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-408 <CHE>
A;Cross-references: EMBL:Z22607; NID:g296856; PIDN:CAA80329.1; PID:g296857
C;Superfamily: inhibin

  Query Match             25.0%;  Score 342.5;  DB 2;  Length 408;
  Best Local Similarity   35.4%;  Pred. No. 7.5e-22;
  Matches   84;  Conservative   31;  Mismatches   73;  Indels   49;  Gaps    7;

Qy      29 WETLEVSSAVKRWVRADSTTNKNKLEVT----VQSHRESCDTLDISVPPGSKN----LPF 80
           ||| :|| || || |         :|||     ::|:     :  |:| || |     |
Db     213 WETFDVSPAVLRWTREKQPNYGLAIEVTHLHQTRTHQGQHVRISRSLPQGSGNWAQLRPL 272

Qy      81 FVVFSNDRSNGTKETRLDLLKEMIGHEQETMLVKTAKNAYQGAGESQEEEGLDGYTAVGP 140
            | | :|                 ||   |:  : || : :
Db     273 LVTFGHDGR---------------GH---TLTRRRAKRSPK------------------- 295

Qy     141 LLARRKRSTGASSHCQKTSLRVNFEDIGWDSWIIAPKEYDAYECKGGCFFPLADDVTPTK 200
                :||   : :|:: || |:| |:||: ||:||  | |: | | | ||||| :  |
Db     296 --HHPQRSRKKNKNCRRHSLYVDFSDVGWNDWIVAPPGYQAFYCHGDCPFPLADHLNSTN 353

Qy     201 HAIVQTLVHLKFPTKVGKACCVPTKLSPISILYKDDMGVPTLKYHYEGMSVAECGCR 257
           ||||||||:    : : |||||||:|| ||:|| |:     || :|: | |  ||||
Db     354 HAIVQTLVN-SVNSSIPKACCVPTELSAISMLYLDEYDKVVLK-NYQEMVVEGCGCR 408


RESULT 4
S58791
bone morphogenetic protein 4 - fallow deer
C;Species: Dama dama (fallow deer)
C;Date: 15-Feb-1996 #sequence_revision 01-Mar-1996 #text_change 16-Feb-1997
C;Accession: S58791
R;Feng, J.Q.; Chen, D.; Esparza, J.; Harris, M.A.; Mundy, G.R.; Harris, S.E.
Biochim. Biophys. Acta 1263, 163-168, 1995
A;Title: Deer antler tissue contains two types of bone morphogenetic protein 4 mRNA t
A;Reference number: S58791; MUID:95367593; PMID:7640308
A;Accession: S58791
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-408 <FEN>
C;Superfamily: inhibin

  Query Match             24.9%;  Score 341;  DB 2;  Length 408;
  Best Local Similarity   35.9%;  Pred. No. 1e-21;
  Matches   85;  Conservative   31;  Mismatches   73;  Indels   48;  Gaps    8;

Qy      29 WETLEVSSAVKRWVRADSTTNKNKLEVT----VQSHRESCDTLDISVPPGSKN----LPF 80
           ||| :|| || || |         :|||     ::|:     :  |:| || :     |
Db     212 WETFDVSPAVLRWTREKQPNYGLAIEVTHLHQTRTHQGQHVRISRSLPQGSGDWAQLRPL 271

Qy      81 FVVFSNDRSNGTKETRLDLLKEMIGHEQETMLVKTAKNAYQGAGESQEEEGLDGYTAVGP 140
            | | :|                 ||         |   :: |  | :           |
Db     272 LVTFGHDGR---------------GH---------ALTRHRRAKRSPKHH---------P 298

Qy     141 LLARRKRSTGASSHCQKTSLRVNFEDIGWDSWIIAPKEYDAYECKGGCFFPLADDVTPTK 200
             ||:|     : :|:: || |:| |:||: ||:||  | |: | | | ||||| :  |
Db     299 QRARKK-----NKNCRRHSLYVDFSDVGWNDWIVAPPGYQAFYCHGDCPFPLADHLNSTN 353

Qy     201 HAIVQTLVHLKFPTKVGKACCVPTKLSPISILYKDDMGVPTLKYHYEGMSVAECGCR 257
           ||||||||:    : : |||||||:|| ||:|| |:     || :|: | |  ||||
Db     354 HAIVQTLVN-SVNSSIPKACCVPTELSAISMLYLDEYDKVVLK-NYQEMVVEGCGCR 408


RESULT 5
I49541
bone morphogenetic protein 4 - mouse
C;Species: Mus musculus (house mouse)
C;Date: 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 16-Jul-1999
C;Accession: I49541; S29523; B34201
R;Feng, J.Q.; Chen, D.; Cooney, A.J.; Tsai, M.; Harris, M.A.; Tsai, S.Y.; Feng, M.; M
J. Biol. Chem. 270, 28364-28373, 1995
A;Title: The mouse bone morphogenetic protein-4(BMP4) gene: Analysis of promoter util
A;Reference number: I49541; MUID:96081880; PMID:7499338
A;Accession: I49541
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-420 <RES>
A;Cross-references: GB:L47480; NID:g994733; PIDN:AAC37698.1; PID:g994734
R;Dickinson, M.E.; van der Meer-de Jong, R.; Hogan, B.L.M.
submitted to the EMBL Data Library, December 1990
A;Description: Nucleotide sequence of the mouse Bone Morphogenetic Protein-4 (BMP-4)
A;Reference number: S29523
A;Accession: S29523
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 13-420 <DIC>
A;Cross-references: EMBL:X56848; NID:g50180; PIDN:CAA40179.1; PID:g50181
R;Dickinson, M.E.; Kobrin, M.S.; Silan, C.M.; Kingsley, D.M.; Justice, M.J.; Miller,
peland, N.G.; Jenkins, N.A.
Genomics 6, 505-520, 1990
A;Title: Chromosomal localization of seven members of the murine TGF-beta superfamily
A;Reference number: A34201; MUID:90228966; PMID:1970330
A;Accession: B34201
A;Status: preliminary; nucleic acid sequence not shown; not compared with conceptual
A;Molecule type: mRNA
```

```
A;Residues: 253-420 <DI2>
C;Genetics:
A;Gene: BMP-4
A;Introns: 11/1; 137/1
C;Superfamily: inhibin

  Query Match             24.7%;  Score 337.5;  DB 2;  Length 420;
  Best Local Similarity   35.0%;  Pred. No. 2.1e-21;
  Matches   83;  Conservative   32;  Mismatches   73;  Indels   49;  Gaps    7;

Qy      29 WETLEVSSAVKRWVRADSTTNKNKLEVT----VQSHRESCDTLDISVPPGSKN----LPF 80
           ||| :|| || || |         :|||     ::|:     :  |:| || :     |
Db     225 WETFDVSPAVLRWTREKQPNYGLAIEVTHLHQTRTHQGQHVRISRSLPQGSGDWAQLRPL 284

Qy      81 FVVFSNDRSNGTKETRLDLLKEMIGHEQETMLVKTAKNAYQGAGESQEEEGLDGYTAVGP 140
             | | :|              ||   |:  : || : :
Db     285 LVTFGHDGR--------------GH---TLTRRRAKRSPK------------------ 307

Qy     141 LLARRKRSTGASSHCQKTSLRVNFEDIGWDSWIIAPKEYDAYECKGGCFFPLADDVTPTK 200
                :||   : :|:: || |:| |:||: ||:|| | |: | | | ||||| :  |
Db     308 --HHPQRSRKKNKNCRRHSLYVDFSDVGWNDWIVAPPGYQAFYCHGDCPFPLADHLNSTN 365

Qy     201 HAIVQTLVHLKFPTKVGKACCVPTKLSPISILYKDDMGVPTLKYHYEGMSVAECGCR 257
           ||||||||:    : : |||||||:|| ||:|| |:     || :|: | |  ||||
Db     366 HAIVQTLVN-SVNSSIPKACCVPTELSAISMLYLDEYDKVVLK-NYQEMVVEGCGCR 420


RESULT 6
B55452
cartilage-derived morphogenetic protein 2 precursor - bovine (fragment)
C;Species: Bos primigenius taurus (cattle)
C;Date: 10-Feb-1995 #sequence_revision 10-Feb-1995 #text_change 26-May-2000
C;Accession: B55452
R;Chang, S.C.; Hoang, B.; Thomas, J.T.; Vukicevic, S.; Luyten, F.P.; Ryba, N.J.P.; Kozak
J. Biol. Chem. 269, 28227-28234, 1994
A;Title: Cartilage-derived morphogenetic proteins. New members of the transforming growt
A;Reference number: A55452; MUID:95050604; PMID:7961761
A;Accession: B55452
A;Status: preliminary; not compared with conceptual translation
A;Molecule type: mRNA
A;Residues: 1-436 <CHA>
A;Cross-references: GB:U13661; NID:g632489; PIDN:AAA61416.1; PID:g632490
C;Superfamily: inhibin

  Query Match             24.1%;  Score 330;  DB 2;  Length 436;
  Best Local Similarity   36.7%;  Pred. No. 9.8e-21;
  Matches   81;  Conservative   23;  Mismatches   67;  Indels   50;  Gaps    7;

Qy      72 PPGSKNLPF------------FVVFSNDRSNGTKETRLDLLKEMIGHEQETMLVKTAKNA 119
           ||  ::| |              ||||  :    ::|    ::| :|   | :
Db     231 PPDLRSLGFGRRVRTPQERALLVVFSRSQ----RKTLFAEMREQLGSATEVV-------- 278

Qy     120 YQGAGESQEEEG---------LDGYTAVG---PLLARRKRSTGASSH----------CQ 156
             | |   |  |           |    |   |   ||:|:  || |           |
Db     279 --GPGGGAEGSGPPPPPPPPPPSGTPDAGLWSPSPGRRRRTAFASRHGKRHGKKSRLRCS 336

Qy     157 KTSLRVNFEDIGWDSWIIAPKEYDAYECKGGCFFPLADDVTPTKHAIVQTLVHLKFPTKV 216
           |  | |||:::||| ||||| ||:|| |:| | |||   : || |||:|||::   |
Db     337 KKPLHVNFKELGWDDWIIAPLEYEAYHCEGVCDFPLRSHLEPTNHAIIQTLMNSMDPGST 396

Qy     217 GKACCVPTKLSPISILYKDDMGVPTLKYHYEGMSVAECGCR 257
           :|||||||:||||||  | |   :   || | |  ||||
Db     397 PPSCCVPTKLTPISILY-IDAGNNVVYNEYEEMVVESCGCR 436


RESULT 7
I50608
bone morphogenetic protein 4 - chicken
C;Species: Gallus gallus (chicken)
C;Date: 13-Sep-1996 #sequence_revision 13-Sep-1996 #text_change 16-Jul-1999
C;Accession: I50608
R;Francis, P.H.; Richardson, M.K.; Brickell, P.M.; Tickle, C.
Development 120, 209-218, 1994
A;Title: Bone morphogenetic proteins and a signalling pathway that controls patternin
A;Reference number: I50607; MUID:94163974; PMID:8119128
A;Accession: I50608
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-405 <FRA>
A;Cross-references: EMBL:X75915; NID:g472929; PIDN:CAA53514.1; PID:g472930
C;Genetics:
A;Gene: Bmp-4
C;Superfamily: inhibin

  Query Match             23.9%;  Score 327.5;  DB 2;  Length 405;
  Best Local Similarity   33.2%;  Pred. No. 1.5e-20;
  Matches   79;  Conservative   36;  Mismatches   70;  Indels   53;  Gaps    7;

Qy      29 WETLEVSSAVKRWVRADSTTNKNKLEVT----VQSHRESCDTLDISVPPG-----SKNLP 79
           ||| :|| || || :     :   :|||     |:|:     :  |:| |     ::  |
Db     212 WETFDVSPAVIRWTKDKQPNHGLVIEVTHLHQAQTHQGKHVRISRSLPQGHGGDWAQLRP 271

Qy      80 FFVVFSNDRSNGTKETRLDLLKEMIGHEQETMLVKTAKNAYQGAGESQEEEGLDGYTAVG 139
             | | :|              ||    | : |: : :  |  : ::
Db     272 LLVTFGHDGR--------------GH----ALTRRARRSPKHHGSRKNKK--------- 303

Qy     140 PLLARRKRSTGASSHCQKTSLRVNFEDIGWDSWIIAPKEYDAYECKGGCFFPLADDVTPT 199
                      :|:: :| |:| |:||: ||:||  | |: | | | |||||  :  |
Db     304 --------------NCRRHALYVDFSDVGWNDWIVAPPGYQAFYCHGDCPFPLADHLNST 349

Qy     200 KHAIVQTLVHLKFPTKVGKACCVPTKLSPISILYKDDMGVPTLKYHYEGMSVAECGCR 257
           ||||||||:    : : |||||||:|| ||:|| |:     || :|: | |  ||||
Db     350 NHAIVQTLVN-SVNSSIPKACCVPTELSAISMLYLDEYDKVVLK-NYQEMVVEGCGCR 405


RESULT 8
JH0801
bone morphogenetic protein 4 - mouse
C;Species: Mus musculus (house mouse)
C;Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 21-Jul-2000
C;Accession: JH0801
R;Kurihara, T.; Kitamura, K.; Takaoka, K.; Nakazato, H.
Biochem. Biophys. Res. Commun. 192, 1049-1056, 1993
A;Title: Murine bone morphogenetic protein-4 gene: existence of multiple promoters an
A;Reference number: JH0801; MUID:93282803; PMID:8507180
A;Accession: JH0801
A;Molecule type: DNA
A;Residues: 1-408 <KUR>
A;Cross-references: GB:D14814; NID:g286069; PIDN:BAA03555.1; PID:g303650
C;Genetics:
A;Introns: 125/1
C;Superfamily: inhibin
C;Keywords: bone

  Query Match             23.9%;  Score 327.5;  DB 2;  Length 408;
  Best Local Similarity   34.6%;  Pred. No. 1.5e-20;
  Matches   82;  Conservative   32;  Mismatches   74;  Indels   49;  Gaps    7;

Qy      29 WETLEVSSAVKRWVRADSTTNKNKLEVT----VQSHRESCDTLDISVPPGSKN----LPF 80
           ||| :|| || || |         :|||     ::|:     :  |:| || :     |
Db     213 WETFDVSPAVLRWTREKQPNYGLAIEVTHLHQTRTHQGQHVRISRSLPQGSGDWAQLRPL 272

Qy      81 FVVFSNDRSNGTKETRLDLLKEMIGHEQETMLVKTAKNAYQGAGESQEEEGLDGYTAVGP 140
             | | :|              ||   |:  : || : :
Db     273 LVTFGHDGR--------------GH---TLTRRRAKRSPK------------------ 295

Qy     141 LLARRKRSTGASSHCQKTSLRVNFEDIGWDSWIIAPKEYDAYECKGGCFFPLADDVTPTK 200
                :||   : :|:: || |:| |:|||: ||:||  | |: | | | ||||| :  |
Db     296 --HHPQRSRKKNKNCRRHSLYVDFSDVGWNDWIVAPPGYQAFYCHGDCPFPLADHLNSTN 353

Qy     201 HAIVQTLVHLKFPTKVGKACCVPTKLSPISILYKDDMGVPTLKYHYEGMSVAECGCR 257
           ||||||||:    : : |||||||:|| ||:|| |:     || :|: | |  || |
Db     354 HAIVQTLVN-SVNSSIPKACCVPTELSAISMLYLDEYDKVVLK-NYQEMVVEGCGSR 408


RESULT 9
```

```
A49147
bone morphogenetic protein 4 - African clawed frog
N;Alternate names: BMP-4; ventralizing factor
C;Species: Xenopus laevis (African clawed frog)
C;Date: 19-Dec-1993 #sequence_revision 18-Nov-1994 #text_change 16-Jul-1999
C;Accession: A49147
R;Dale, L.; Howes, G.; Price, B.M.; Smith, J.C.
Development 115, 573-585, 1992
A;Title: Bone morphogenetic protein 4: a ventralizing factor in early Xenopus developmen
A;Reference number: A49147; MUID:93048819; PMID:1425340
A;Accession: A49147
A;Status: preliminary
A;Molecule type: nucleic acid
A;Residues: 1-400 <DAL>
A;Cross-references: GB:X64538; GB:S46999; NID:g64589; PIDN:CAA45836.1; PID:g64590
A;Experimental source: XTC cells
A;Note: sequence extracted from NCBI backbone (NCBIN:117127, NCBIP:117128)
C;Superfamily: inhibin

  Query Match             23.9%;  Score 326.5;  DB 2;  Length 400;
  Best Local Similarity   32.8%;  Pred. No. 1.8e-20;
  Matches   78;  Conservative   35;  Mismatches   72;  Indels   53;  Gaps    7;

Qy      29 WETLEVSSAVKRWVRADSTTNKNKLEV----TVQSHRESCDTLDISVPP-----GSKNLP 79
           ||: :|| |: || |     :   :||      ::|:     :  |: |      |:  |
Db     207 WESFDVSPAIMRWTRDKQINHGLAIEVIHLNQTKTHQGKHVRISRSLLPQEDADWSQMRP 266

Qy      80 FFVVFSNDRSNGTKETRLDLLKEMIGHEQETMLVKTAKNAYQGAGESQEEEGLDGYTAVG 139
             : ||:|              ||    | : :| :
Db     267 LLITFSHDGR--------------GH----ALTRRSKRS------------------- 287

Qy     140 PLLARRKRSTGASSHCQKTSLRVNFEDIGWDSWIIAPKEYDAYECKGGCFPPLADDVTPT 199
               :::|    : ||:: || |:| |:||: ||:||  | |: | | | |||||  :  |
Db     288 ---PKQQRPRKKNKHCRRHSLYVDFSDVGWNDWIVAPPGYQAFYCHGDCPFPLADHLNST 344

Qy     200 KHAIVQTLVHLKFPTKVGKACCVPTKLSPISILYKDDMGVPTLKYHYEGMSVAECGCR 257
           ||||||||:      : |||||||:|| ||:|| |:     || :|: | |  ||||
Db     345 NHAIVQTLVN-SVNASIPKACCVPTELSAISMLYLDEYDKVVLK-NYQEMVVEGCGCR 400


RESULT 10
A26158
decapentaplegic protein precursor - fruit fly (Drosophila melanogaster)
C;Species: Drosophila melanogaster
C;Date: 24-Jun-1987 #sequence_revision 24-Jun-1987 #text_change 16-Apr-1999
C;Accession: A26158
R;Padgett, R.W.; St. Johnston, R.D.; Gelbart, W.M.
Nature 325, 81-84, 1987
A;Title: A transcript from a Drosophila pattern gene predicts a protein homologous to th
A;Reference number: A26158; MUID:87090408; PMID:3467201
A;Accession: A26158
A;Molecule type: mRNA
A;Residues: 1-588 <PAD>
A;Cross-references: GB:M30116; NID:g157291; PID:g157292
C;Genetics:
A;Gene: FlyBase:dpp
A;Cross-references: FlyBase:FBgn0000490
C;Keywords: glycoprotein
F;1-15/Domain: signal sequence #status predicted <SIG>
F;120,342,377,529/Binding site: carbohydrate (Asn) (covalent) #status predicted

  Query Match             23.8%;  Score 325;  DB 2;  Length 588;
  Best Local Similarity   30.5%;  Pred. No. 3.9e-20;
  Matches   92;  Conservative   45;  Mismatches  115;  Indels   50;  Gaps    9;

Qy       4 EDSETWDQATGTKTFLVSQDIR----DEGWETLEV----SSAVKRWVRADSTTNKNKLEV 55
           :||:  |:      | :  |::    ||  :  |:     :  :: | : |: |: : :|
Db     289 KDSKIDDRFPHHHRFRLHFDVKSIPADEKLKAAELQLTRDALSQQVVASRSSANRTRYQV 348

Qy      56 T--------VQSHRE-------------SCDTLDISVPPG--------SKNLPFFVVFS 85
                  |:  ||              | ||: : | |           :|    |
Db     349 LVYDITRVGVRGQREPSYLLLDTKTVRLNSTDTVSLDVQPAVDRWLASPQRNYGLLVEVR 408

Qy      86 NDRS-------NGTKETRLDLLKEMIGHEQETMLVKTAKNAYQGAGESQEEEGLDGYTAV 138
             ||       :       |  |  |:|  :  |    :: |   ::  | :|
Db     409 TVRSLKPAPHHHVRLRRSADEAHERWQHKQPLLFTYTDDGRHK-ARSIRDVSGGEGGGKG 467

Qy     139 GPLLARRKRSTGASSH---CQKTSLRVNFEDIGWDSWIIAPKEYDAYECKGGCFPPLADD 195
           |       :| |   :|    |:: || |:| |:||| ||:||  |||| | | | |||||
Db     468 GRNKRHARRPTRRKNHDDTCRRHSLYVDFSDVGWDDWIVAPLGYDAYYCHGKCPFPLADH 527

Qy     196 VTPTKHAIVQTLVHLKFPTKVGKACCVPTKLSPISILYKDDMGVPTLKYHYEGMSVAECG 255
           | ||:|||||:   | || |||||||:|  :::|| :|     || :|: |:|  ||
Db     528 FNSTNHAVVQTLVNNMNPGKVPKACCVPTQLDSVAMLYLNDQSTVVLK-NYQEMTVVGCG 586

Qy     256 CR 257
           ||
Db     587 CR 588


RESULT 11
JH0688
bone morphogenetic protein 2II precursor - African clawed frog
C;Species: Xenopus laevis (African clawed frog)
C;Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 16-Jul-1999
C;Accession: JH0688
R;Nishimatsu, S.; Suzuki, A.; Shoda, A.; Murakami, K.; Ueno, N.
Biochem. Biophys. Res. Commun. 186, 1487-1495, 1992
A;Title: Genes for bone morphogenetic proteins are differentially transcribed in earl
A;Reference number: JH0687; MUID:92378616; PMID:1510675
A;Accession: JH0688
A;Molecule type: mRNA
A;Residues: 1-398 <NIS>
A;Cross-references: GB:X63425; NID:g64583; PIDN:CAA45019.1; PID:g64584
A;Experimental source: oocyte
C;Superfamily: inhibin
C;Keywords: glycoprotein
F;285-398/Product: bone morphogenetic protein 2II #status predicted <MAT>
F;137,202,237,340/Binding site: carbohydrate (Asn) (covalent) #status predicted

  Query Match             23.3%;  Score 319;  DB 2;  Length 398;
  Best Local Similarity   34.0%;  Pred. No. 7.8e-20;
  Matches   83;  Conservative   29;  Mismatches   66;  Indels   66;  Gaps    9;

Qy      29 WETLEVSSAVKRWVRADSTTNKNKLEVTVQSHRESCDTLDISVPPG----SKNL------ 78
           ||: :|: |: ||:      :   :|||   | ::    | :||      |::|
Db     206 WESFDVTPAITRWIAHKQPNHGFVVEVT---HLDN----DTNVPKRHVRISRSLTLDKGH 258

Qy      79 -----PFFVVFSNDRSNGTKETRLDLLKEMIGHEQETMLVKTAKNAYQGAGESQEEEGLD 133
                |  | ||:|                  ||     | |  |
Db     259 WPRIRPLLVTFSHDGK--------------GH----ALHKRQKRQ-------------- 285

Qy     134 GYTAVGPLLARRKRSTGASSHCQKTSLRVNFEDIGWDSWIIAPKEYDAYECKGGCFPPLA 193
                    || |:     | |::  | |:| |:|| ||:||  | |: | | | ||||
Db     286 ---------ARHKQRKRLKSSCRRHPLYVDFSDVGWNDWIVAPPGYHAFYCHGECPFPLA 336

Qy     194 DDVTPTKHAIVQTLVHLKFPTKVGKACCVPTKLSPISILYKDDMGVPTLKYHYEGMSVAE 253
           | :  | ||||||||:     | : |||||||:|| ||:|| |:     || :|: | |
Db     337 DHLNSTNHAIVQTLVN-SVNTNIPKACCVPTELSAISMLYLDENEKVVLK-NYQDMVVEG 394

Qy     254 CGCR 257
           ||||
Db     395 CGCR 398


RESULT 12
JH0687
bone morphogenetic protein 2I precursor - African clawed frog
C;Species: Xenopus laevis (African clawed frog)
C;Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 16-Jul-1999
C;Accession: JH0687; S16244
R;Nishimatsu, S.; Suzuki, A.; Shoda, A.; Murakami, K.; Ueno, N.
Biochem. Biophys. Res. Commun. 186, 1487-1495, 1992
A;Title: Genes for bone morphogenetic proteins are differentially transcribed in earl
A;Reference number: JH0687; MUID:92378616; PMID:1510675
A;Accession: JH0687
```

```
A;Molecule type: mRNA
A;Residues: 1-398 <NIS>
A;Cross-references: GB:X63424; NID:g64585; PIDN:CAA45018.1; PID:g64586
A;Experimental source: oocyte
R;Plessow, S.; Koester, M.; Knoechel, W.
Biochim. Biophys. Acta 1089, 280-282, 1991
A;Title: cDNA sequence of Xenopus laevis bone morphogenetic protein 2 (BMP-2).
A;Reference number: S16244; MUID:91274367; PMID:2054389
A;Accession: S16244
A;Molecule type: mRNA
A;Residues: 1-6,'S',8-15,'V',17-232,'N',234-398 <PLE>
A;Cross-references: EMBL:X55031; NID:g64581; PIDN:CAA38850.1; PID:g64582
C;Superfamily: inhibin
C;Keywords: dimer; glycoprotein
F;285-398/Product: bone morphogenetic protein 2I #status predicted <MAT>
F;137,202,340/Binding site: carbohydrate (Asn) (covalent) #status predicted

  Query Match             23.2%;  Score 318;  DB 2;  Length 398;
  Best Local Similarity   33.8%;  Pred. No. 9.5e-20;
  Matches   80;  Conservative   29;  Mismatches   76;  Indels   52;  Gaps    7;

Qy      29 WETLEVSSAVKRWVRADSTTNKNKLEVTVQSHRESCDTLDI----SVPPGSKN----LPF 80
           ||: :|: |: ||:      :   :|||   : ::     :    |: |   |     |
Db     206 WESFDVTPAIARWIAHKQPNHGFVVEVTHLDNDKNVPKKHVRISRSLTPDKDNWPQIRPL 265

Qy      81 FVVFSNDRSNGTKETRLDLLKEMIGHEQETMLVKTAKNAYQGAGESQEEEGLDGYTAVGP 140
             | ||:|                 ||     | |  |
Db     266 LVTFSHDGK--------------GH----ALHKRQKRQ---------------------- 285

Qy     141 LLARRKRSTGASSHCQKTSLRVNFEDIGWDSWIIAPKEYDAYECKGGCFFPLADDVTPTK 200
             || |:     | |::  | |:| |:||: ||:||  | |: | | | ||||| :  |
Db     286 --ARHKQRKRLKSSCRRHPLYVDFSDVGWNDWIVAPPGYHAFYCHGECPFPLADHLNSTN 343

Qy     201 HAIVQTLVHLKFPTKVGKACCVPTKLSPISILYKDDMGVPTLKYHYEGMSVAECGCR 257
           ||||||||:    | : |||||||:|| ||:|| |:     || :|: | |  ||||
Db     344 HAIVQTLVN-SVNTNIPKACCVPTELSAISMLYLDENEKVVLK-NYQDMVVEGCGCR 398


RESULT 13
JH0689
bone morphogenetic protein 4 precursor - African clawed frog
C;Species: Xenopus laevis (African clawed frog)
C;Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 16-Jul-1999
C;Accession: JH0689
R;Nishimatsu, S.; Suzuki, A.; Shoda, A.; Murakami, K.; Ueno, N.
Biochem. Biophys. Res. Commun. 186, 1487-1495, 1992
A;Title: Genes for bone morphogenetic proteins are differentially transcribed in early a
A;Reference number: JH0687; MUID:92378616; PMID:1510675
A;Accession: JH0689
A;Molecule type: mRNA
A;Residues: 1-401 <NIS>
A;Cross-references: GB:X63426; NID:g64587; PIDN:CAA45020.1; PID:g64588
A;Experimental source: oocyte
C;Superfamily: inhibin
C;Keywords: glycoprotein
F;288-401/Product: bone morphogenetic protein 4 #status predicted <MAT>
F;141,204,238,343,358/Binding site: carbohydrate (Asn) (covalent) #status predicted

  Query Match             23.1%;  Score 316.5;  DB 2;  Length 401;
  Best Local Similarity   33.2%;  Pred. No. 1.3e-19;
  Matches   79;  Conservative   31;  Mismatches   75;  Indels   53;  Gaps    7;

Qy      29 WETLEVSSAVKRWVRADSTTNKNKLEV-------TVQSH--RESCDTLDISVPPGSKNLP 79
           ||: :|| |: ||      :   :||       | |    | |   |      |:  |
Db     208 WESFDVSPAIMRWTLDKQINHGLAIEVIHLNQTKTYQGKHVRISRSLLPQKDADWSQMRP 267

Qy      80 FFVVFSNDRSNGTKETRLDLLKEMIGHEQETMLVKTAKNAYQGAGESQEEEGLDGYTAVG 139
             : ||:|                 ||     | : :| :
Db     268 LLITFSHDGR--------------GH----ALTRRSKRS------------------- 288

Qy     140 PLLARRKRSTGASSHCQKTSLRVNFEDIGWDSWIIAPKEYDAYECKGGCFFPLADDVTPT 199
              :::|    : ||:: || |:| |:||: ||:||  | |: | | | |||||  :  |
Db     289 ---PKQQRPRKKNKHCRRHSLYVDFSDVGWNDWIVAPPGYQAFYCHGDCPFPLADHLNST 345

Qy     200 KHAIVQTLVHLKFPTKVGKACCVPTKLSPISILYKDDMGVPTLKYHYEGMSVAECGCR 257
           ||||||||:    : : |||||||:|| ||:|| |:     || :|: | |  ||||
Db     346 NHAIVQTLVN-SVNSSIPKACCVPTELSAISMLYLDEYDKVVLK-NYQEMVVEGCGCR 401


RESULT 14
JQ1184
osteogenic protein 1 precursor - mouse
C;Species: Mus musculus (house mouse)
C;Date: 31-Dec-1991 #sequence_revision 31-Dec-1991 #text_change 06-Dec-1996
C;Accession: JQ1184; PQ0224
R;Oezkaynak, E.; Schnegelsberg, P.N.J.; Oppermann, H.
Biochem. Biophys. Res. Commun. 179, 116-123, 1991
A;Title: Murine osteogenic protein (OP-1): high levels of mRNA in kidney.
A;Reference number: JQ1184; MUID:91354237; PMID:1715687
A;Accession: JQ1184
A;Molecule type: DNA
A;Residues: 1-120 <OEZ>
A;Cross-references: EMBL:X56906
A;Accession: PQ0224
A;Molecule type: mRNA
A;Residues: 54-430 <OE2>
C;Comment: This protein induces bone formation.
C;Superfamily: inhibin
C;Keywords: glycoprotein
F;1-29/Domain: signal sequence #status predicted <SIG>
F;30-290/Domain: propeptide #status predicted <PRO>
F;291-430/Product: osteogenic protein-1 #status predicted <MAT>
F;186,301,320,371/Binding site: carbohydrate (Asn) (covalent) #status predicted

  Query Match             23.0%;  Score 314.5;  DB 2;  Length 430;
  Best Local Similarity   32.6%;  Pred. No. 2.1e-19;
  Matches   90;  Conservative   33;  Mismatches   98;  Indels   55;  Gaps   11;

Qy       5 DSETW--------DQATGTKTFLVSQDIR-----DEGWETLEVSSAVKRWVRADSTTNKN 51
           |:||:        : :| :: |  | |      :|||   ::::    ||       ::
Db     185 DNETFQITVYQVLQEHSGRESDLFLLDSRTIWASEEGWLVFDITATSNHWV----VNPRH 240

Qy      52 KLEVTVQSHRESCDTLD-ISVPP---------GSKN-LPFFVVFSNDRSNGTKETRLDLL 100
           | : :     | :|||  |: |          | :|  || | |          |  |  :
Db     241 NLGLQL-----SVETLDGQSINPKLAGLIGRHGPQNKQPFMVAFFK-----ATEVHLRSI 290

Qy     101 KEMIGHEQETMLVKTAKNAYQGAGESQEEEGLDGYTAVGPLLARRKRSTGASSHCQKTSL 160
           :   | ::     || ||         ||         |:         |:     |:|  |
Db     291 RSTGGKQRSQNRSKTPKN--------QE--------ALRMASVAENSSSDQRQACKKHEL 334

Qy     161 RVNFEDIGWDSWIIAPKEYDAYECKGGCFFPLADDVTPTKHAIVQTLVHLKFPTKVGKAC 220
           |:| |:||  ||||||: | || |:| | |||   :  | |||||||||   |  | | |
Db     335 YVSFRDLGWQDWIIAPEGYAAYYCEGECAFPLNSYMNATNHAIVQTLVHFINPDTVPKPC 394

Qy     221 CVPTKLSPISILYKDDMGVPTLKYHYEGMSVAECGC 256
           | ||:|: ||:|| ||     ||  |  | |  |||
Db     395 CAPTQLNAISVLYFDDSSNVILK-KYRNMVVRACGC 429


RESULT 15
S37073
bone morphogenetic protein 2 - rat
C;Species: Rattus norvegicus (Norway rat)
C;Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 16-Jul-1999
C;Accession: S37073
R;Feng, J.Q.; Chen, D.; Feng, M.; Harris, M.A.; Mundy, G.R.; Harris, S.E.
submitted to the EMBL Data Library, September 1993
A;Description: cDNA sequence of fetal rat calvarial osteoblast bone morphogenetic pro
A;Reference number: S37073
A;Accession: S37073
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-393 <FEN>
A;Cross-references: EMBL:Z25868; NID:g397950; PIDN:CAA81088.1; PID:g397951
C;Superfamily: inhibin
```

```
  Query Match             23.0%;  Score 314;  DB 2;  Length 393;
  Best Local Similarity   33.1%;  Pred. No. 2.1e-19;
  Matches   86;  Conservative   30;  Mismatches   68;  Indels   76;  Gaps   10;

Qy      17 TFLVSQDIRDEGWETLEVSSAVKRWVRADSTTNKNKLEVTVQSHRESCDTLDISVPPG-- 74
           | ||:|:     ||: :|: || ||     | :   :||   :| |          ||
Db     191 TRLVTQNTSQ--WESFDVTPAVMRWTAQGHTNHGFVVEV---AHLEE--------KPGVS 237

Qy      75 ------SKNL-----------PFFVVFSNDRSNGTKETRLDLLKEMIGHEQETMLVKTAK 117
                 |::|           |  | | :|                ||       : ||
Db     238 KRHVRISRSLHQDEHSWSQVRPLLVTFGHDGK--------------GHPLHKREKRQAK 282

Qy     118 NAYQGAGESQEEEGLDGYTAVGPLLARRKRSTGASSHCQKTSLRVNFEDIGWDSWIIAPK 177
           :                      :|||     | |::  | |:| |:||: ||:||
Db     283 H----------------------KQRKR---LKSSCKRHPLYVDFSDVGWNDWIVAPP 315

Qy     178 EYDAYECKGGCFFPLADDVTPTKHAIVQTLVHLKFPTKVGKACCVPTKLSPISILYKDDM 237
           | |: | | | ||||| :  | ||||||||:    :|: |||||||:|| ||:|| |:
Db     316 GYHAFYCHGECPFPLADHLNSTNHAIVQTLVN-SVNSKIPKACCVPTELSAISMLYLDEN 374

Qy     238 GVPTLKYHYEGMSVAECGCR 257
               || :|: | |  ||||
Db     375 EKVVLK-NYQDMVVEGCGCR 393


Search completed: April  2, 2003, 14:52:31
Job time : 45 secs
```

OM protein - protein search, using sw model

Run on: April 2, 2003, 14:51:46 ; Search time 36 Seconds
(without alignments)
436.443 Million cell updates/sec

Title: US-10-002-278-9
Perfect score: 1368
Sequence: 1 DVLEDSETWDQATGTKTFLV.....GVPTLKYHYEGMSVAECGCR 257

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 248812 seqs, 61136040 residues

Total number of hits satisfying chosen parameters: 248812

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published_Applications_AA:*
1: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep:*
2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep:*
3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep:*
4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep:*
5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep:*
6: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep:*
7: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep:*
8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep:*
9: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep:*
10: /cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB.pep:*
11: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep:*
12: /cgn2_6/ptodata/2/pubpaa/US10_PUBCOMB.pep:*
13: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep:*
14: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	1368	100.0	257	12	US-10-002-278-9	Sequence 9, Appli
2	808.5	59.1	427	12	US-10-002-278-2	Sequence 2, Appli
3	516.5	37.8	425	9	US-09-813-398-32	Sequence 32, Appl
4	361	26.4	419	10	US-09-784-911-2	Sequence 2, Appli
5	354	25.9	437	10	US-09-784-911-4	Sequence 4, Appli
6	350	25.6	427	10	US-09-784-911-6	Sequence 6, Appli
7	349.5	25.5	321	9	US-09-945-182-26	Sequence 26, Appl
8	345	25.2	408	9	US-09-804-625-6	Sequence 6, Appli
9	345	25.2	408	10	US-09-749-728B-69	Sequence 69, Appl
10	345	25.2	409	9	US-09-813-398-27	Sequence 27, Appl
11	337.5	24.7	263	9	US-09-945-182-32	Sequence 32, Appl
12	334.5	24.5	411	9	US-09-945-182-28	Sequence 28, Appl
13	330	24.1	436	10	US-09-730-772-14	Sequence 14, Appl
14	330	24.1	436	10	US-09-735-849-14	Sequence 14, Appl
15	322	23.5	451	10	US-09-784-911-10	Sequence 10, Appl
16	320	23.4	433	10	US-09-784-911-8	Sequence 8, Appli
17	313	22.9	433	10	US-09-784-911-12	Sequence 12, Appl
18	312.5	22.8	455	9	US-09-813-398-28	Sequence 28, Appl
19	312	22.8	435	10	US-09-784-911-14	Sequence 14, Appl
20	311	22.7	396	9	US-10-044-716-2	Sequence 2, Appli
21	311	22.7	396	9	US-09-804-625-4	Sequence 4, Appli
22	311	22.7	396	10	US-09-952-360-2	Sequence 2, Appli
23	311	22.7	397	9	US-09-813-398-24	Sequence 24, Appl
24	310	22.7	118	12	US-10-115-406-15	Sequence 15, Appl
25	310	22.7	143	12	US-10-002-278-4	Sequence 4, Appli
26	308	22.5	502	9	US-09-813-398-37	Sequence 37, Appl
27	307	22.4	501	8	US-08-981-490B-1	Sequence 1, Appli
28	306	22.4	501	10	US-09-730-772-13	Sequence 13, Appl
29	306	22.4	501	10	US-09-735-849-13	Sequence 13, Appl
30	305.5	22.3	431	8	US-08-822-186-2	Sequence 2, Appli
31	305.5	22.3	431	8	US-08-937-755-2	Sequence 2, Appli
32	305.5	22.3	431	9	US-09-982-543A-10	Sequence 10, Appl
33	305.5	22.3	431	9	US-10-062-370-5	Sequence 5, Appli
34	305.5	22.3	431	9	US-09-039-107-2	Sequence 2, Appli
35	305.5	22.3	431	9	US-09-798-518-1	Sequence 1, Appli
36	305.5	22.3	431	10	US-09-045-331-2	Sequence 2, Appli
37	305.5	22.3	431	10	US-09-828-607-2	Sequence 2, Appli
38	305.5	22.3	432	9	US-10-270-377-3	Sequence 3, Appli
39	305.5	22.3	432	10	US-09-361-741-3	Sequence 3, Appli
40	305.5	22.3	495	9	US-09-880-708-10	Sequence 10, Appl
41	304	22.2	294	9	US-09-945-182-2	Sequence 2, Appli
42	304	22.2	388	9	US-09-945-182-34	Sequence 34, Appl
43	302.5	22.1	119	8	US-08-945-459A-1	Sequence 1, Appli
44	302.5	22.1	119	9	US-09-068-253-2	Sequence 2, Appli
45	302.5	22.1	120	9	US-09-945-182-4	Sequence 4, Appli

ALIGNMENTS

RESULT 1
US-10-002-278-9
; Sequence 9, Application US/10002278
; Patent No. US20020132334A1
; GENERAL INFORMATION:
; APPLICANT: Jessell, Thomas M.
; APPLICANT: Basler, Konard
; APPLICANT: Yamada, Toshiya
; TITLE OF INVENTION: CLONING, EXPRESSION AND USES OF DORSALIN-1
; FILE REFERENCE: 0575/40314-A
; CURRENT APPLICATION NUMBER: US/10/002,278
; CURRENT FILING DATE: 2001-11-02
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 9
; LENGTH: 257
; TYPE: PRT
; ORGANISM: Mouse
US-10-002-278-9

Query Match 100.0%; Score 1368; DB 12; Length 257;
Best Local Similarity 100.0%; Pred. No. 2.2e-128;
Matches 257; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy       1 DVLEDSETWDQATGTKTFLVSQDIRDEGWETLEVSSAVKRWVRADSTTNKNKLEVTVQSH 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 DVLEDSETWDQATGTKTFLVSQDIRDEGWETLEVSSAVKRWVRADSTTNKNKLEVTVQSH 60

Qy      61 RESCDTLDISVPPGSKNLPFFVVFSNDRSNGTKETRLDLLKEMIGHEQETMLVKTAKNAY 120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      61 RESCDTLDISVPPGSKNLPFFVVFSNDRSNGTKETRLDLLKEMIGHEQETMLVKTAKNAY 120

Qy     121 QGAGESQEEEGLDGYTAVGPLLARRKRSTGASSHCQKTSLRVNFEDIGWDSWIIAPKEYD 180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     121 QGAGESQEEEGLDGYTAVGPLLARRKRSTGASSHCQKTSLRVNFEDIGWDSWIIAPKEYD 180

Qy     181 AYECKGGCFFPLADDVTPTKHAIVQTLVHLKFPTKVGKACCVPTKLSPISILYKDDMGVP 240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     181 AYECKGGCFFPLADDVTPTKHAIVQTLVHLKFPTKVGKACCVPTKLSPISILYKDDMGVP 240

Qy     241 TLKYHYEGMSVAECGCR 257
           |||||||||||||||||
```

```
Db    241 TLKYHYEGMSVAECGCR 257


RESULT 2
US-10-002-278-2
; Sequence 2, Application US/10002278
; Patent No. US20020132334A1
; GENERAL INFORMATION:
;  APPLICANT: Jessell, Thomas M.
;  APPLICANT:  Basler, Konard
;  APPLICANT:  Yamada, Toshiya
;  TITLE OF INVENTION: CLONING, EXPRESSION AND USES OF DORSALIN-1
;  FILE REFERENCE: 0575/40314-A
;  CURRENT APPLICATION NUMBER: US/10/002,278
;  CURRENT FILING DATE:  2001-11-02
;  NUMBER OF SEQ ID NOS: 18
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
;   LENGTH: 427
;   TYPE: PRT
;   ORGANISM: Chick
US-10-002-278-2

  Query Match            59.1%;  Score 808.5;  DB 12;  Length 427;
  Best Local Similarity   61.9%;  Pred. No. 2.4e-72;
  Matches  161;  Conservative   31;  Mismatches   59;  Indels    9;  Gaps    5;

Qy      1 DVLEDSETWDQATGTKTFLVSQDIRDEGWETLEVSSAVKRWVRADSTTNKNKLEVTVQSH 60
          ||| | : |:    ||: ||| |:| ||| ||||||||||:||    |||||| ::|
Db    174 DVL-DGDHWENKESTKSLLVSHSIQDCGWEMFEVSSAVKRWVKADKMKTKNKLEVVIESK 232

Qy     61 RES---CDTLDISVPPGSKNLPFFVVFSNDRSNGTKETRLDLLKEMIGHEQETMLVKTAK 117
            |   | |||:|   :|||| :|||||||||||||:::: |:||| ||||::| |  |
Db    233 DLSGFPCGKLDITVTHDTKNLPLLIVFSNDRSNGTKETKVE-LREMIVHEQESVLNKLGK 291

Qy    118 NAYQGAGESQEEEGLDGYTAVGPLLARRKRSTGASSHCQKTSLRVNFEDIGWDSWIIAPK 177
          |      | :||: :          :| ||| || :||::||| |||::|||||||||||
Db    292 NDSSSEEEQREEKAI---ARPRQHSSRSKRSIGA-NHCRRTSLHVNFKEIGWDSWIIAPK 347

Qy    178 EYDAYECKGGCFFPLADDVTPTKHAIVQTLVHLKFPTKVGKACCVPTKLSPISILYKDDM 237
          :|:|:|||||||||| |:|||||||||||||||: | |  |||||||||  ||||||||
Db    348 DYEAFECKGGCFFPLTDNVTPTKHAIVQTLVHLQNPKKASKACCVPTKLDAISILYKDDA 407

Qy    238 GVPTLKYHYEGMSVAECGCR 257
          ||||| |:|||| |||||||
Db    408 GVPTLIYNYEGMKVAECGCR 427


RESULT 3
US-09-813-398-32
; Sequence 32, Application US/09813398
; Patent No. US20020169292A1
; GENERAL INFORMATION:
;  APPLICANT: Bruce D. Weintraub
;  APPLICANT:  Mariusz W. Szkudlinski
;  APPLICANT:  University of Maryland
;  TITLE OF INVENTION: CYSTINE KNOT GROWTH FACTOR MUTANTS
;  FILE REFERENCE: UOFMD.003C1
;  CURRENT APPLICATION NUMBER: US/09/813,398
;  CURRENT FILING DATE:  2001-03-20
;  PRIOR APPLICATION NUMBER: PCT/US99/05908
;  PRIOR FILING DATE: 1999-03-19
;  PRIOR APPLICATION NUMBER: PCT/US98/19772
;  PRIOR FILING DATE: 1998-09-22
;  NUMBER OF SEQ ID NOS: 41
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 32
;   LENGTH: 425
;   TYPE: PRT
;   ORGANISM: HOMO SAPIEN
US-09-813-398-32

  Query Match            37.8%;  Score 516.5;  DB 9;  Length 425;
  Best Local Similarity   42.8%;  Pred. No. 2.9e-43;
  Matches  119;  Conservative   48;  Mismatches   68;  Indels   43;  Gaps   11;

Qy      1 DVLEDSETWDQATGTKTFLVSQDI--RDEGWETLEVSSAVKRWVRADSTTNKNKLEVTVQ 58
          :||| |:  ::        ||| :|   :  ||| :|: |::|| :: |:|  ::||| ::
Db    170 EVLE-SKGDNEGERNMLVLVSGEIYGTNSEWETFDVTDAIRRWQKSGSST--HQLEVHIE 226

Qy     59 S-HRESCDT----LDISVPPGSKNLPFFVVFSNDRSNGTKETRLDLLKEMIGHEQETMLV 113
          | | |: |     |:|     :|: |  :|||:|:|  : : | : | ||| |||   |
Db    227 SKHDEAEDASSGRLEIDTSAQNKHNPLLIVFSDDQS--SDKERKEELNEMISHEQLPEL- 283

Qy    114 KTAKNAYQGAGESQEEEGLDGYTAVGP----LLARRK----------RSTGASSHCQKTS 159
                      :  ||| ::: ||    ||  |           |     ::|::|
Db    284 -------------DNLGLDSFSS-GPGEEALLQMRSNIIYDSTARIRRNAKGNYCKRTP 328

Qy    160 LRVNFEDIGWDSWIIAPKEYDAYECKGGCFFPLADDVTPTKHAIVQTLVHLKFPTKVGKA 219
          | ::|::||||||||||  |:||||:| | :|||: :||||||:| |||||   |  ||
Db    329 LYIDFKEIGWDSWIIAPPGYEAYECRGVCNYPLAEHLTPTKHAIIQALVHLKNSQKASKA 388

Qy    220 CCVPTKLSPISILYKDDMGVPTLKYHYEGMSVAECGCR 257
          ||||||| ||||||  | || | |: ||||:|:|||||
Db    389 CCVPTKLEPISILYL-DKGVVTYKFKYEGMAVSECGCR 425


RESULT 4
US-09-784-911-2
; Sequence 2, Application US/09784911
; Patent No. US20020072115A1
; GENERAL INFORMATION:
;  APPLICANT: Harrison, Leonard C.
;  APPLICANT:  Jiang, Fang-Xu
;  APPLICANT:  Stanley, Edouard Guy
;  APPLICANT:  Gonez, Leonel Jorge
;  TITLE OF INVENTION: Pancreatic islet cell growth factors
;  FILE REFERENCE: Davies Collison Cave
;  CURRENT APPLICATION NUMBER: US/09/784,911
;  CURRENT FILING DATE:  2001-09-17
;  NUMBER OF SEQ ID NOS: 30
;  SOFTWARE: PatentIn version 2.1
; SEQ ID NO 2
;   LENGTH: 419
;   TYPE: PRT
;   ORGANISM: mouse
;   FEATURE:
;   NAME/KEY: UNSURE
;   LOCATION: (186)
;   OTHER INFORMATION: Xaa at position 186 is Tyr or His
US-09-784-911-2

  Query Match            26.4%;  Score 361;  DB 10;  Length 419;
  Best Local Similarity   37.1%;  Pred. No. 8.6e-28;
  Matches   88;  Conservative   32;  Mismatches   75;  Indels   42;  Gaps    8;

Qy     29 WETLEVSSAVKRWVRADSTTNKNKLEVT----VQSHRESCDTLDISVPPGSKN----LPF 80
          ||| :|| || || |          :|||     ::|:     :  |:| || :     |
Db    217 WETFDVSPAVLRWTREKQPNYGLAIEVTHLHQTRTHQGQHVRISRSLPQGSGDWAQLRPL 276

Qy     81 FVVFSNDRSNGTKETRLDLLKEMIGHEQETMLVKTAKNAYQGAGESQEEEGLDGYTAVGP 140
          | | :|                ||       | : | :    ||||  :       |
Db    277 LVTFGHDGR--------------GH------TLTRRRAKRSPKTRQEEEYM-------P 308

Qy    141 LLARRKRSTGASSHCQKTSLRVNFEDIGWDSWIIAPKEYDAYECKGGCFFPLADDVTPTK 200
          :  |||:    : :|:: || |:| |:||: ||:||  | |: | | | ||||| :  |
Db    309 METRRKK----NKNCRRHSLYVDFSDVGWNDWIVAPPGYQAFYCHGDCPFPLADHLNSTN 364

Qy    201 HAIVQTLVHLKFPTKVGKACCVPTKLSPISILYKDDMGVPTLKYHYEGMSVAECGCR 257
          ||||||||:    : : |||||||:|| ||:|| |:     || :|: | |  ||||
Db    365 HAIVQTLVN-SVNSSIPKACCVPTELSAISMLYLDEYDKVVLK-NYQEMVVEGCGCR 419


RESULT 5
US-09-784-911-4
; Sequence 4, Application US/09784911
```

```
; Patent No. US20020072115A1
; GENERAL INFORMATION:
;  APPLICANT: Harrison, Leonard C.
;  APPLICANT:  Jiang, Fang-Xu
;  APPLICANT:  Stanley, Edouard Guy
;  APPLICANT:  Gonez, Leonel Jorge
;  TITLE OF INVENTION: Pancreatic islet cell growth factors
;  FILE REFERENCE: Davies Collison Cave
;  CURRENT APPLICATION NUMBER: US/09/784,911
;  CURRENT FILING DATE:  2001-09-17
;  NUMBER OF SEQ ID NOS: 30
;  SOFTWARE: PatentIn version 2.1
; SEQ ID NO 4
;   LENGTH: 437
;   TYPE: PRT
;   ORGANISM: mouse
;   FEATURE:
;   NAME/KEY: UNSURE
;   LOCATION: (186)
;   OTHER INFORMATION: Xaa at position 186 is Tyr or His
US-09-784-911-4

  Query Match             25.9%;  Score 354;  DB 10;  Length 437;
  Best Local Similarity   37.7%;  Pred. No. 4.6e-27;
  Matches   90;  Conservative   26;  Mismatches   93;  Indels   30;  Gaps    6;

Qy      29 WETLEVSSAVKRWVRADSTTNKNKLEVT----VQSHRESCDTLDISVPPGSKN----LPF 80
           ||| :|| || || |          :|||     ::|:     :  |:| || :     |
Db     217 WETFDVSPAVLRWTREKQPNYGLAIEVTHLHQTRTHQGQHVRISRSLPQGSGDWAQLRPL 276

Qy      81 FVVFSNDRSNGTKETRLDLLKEMIGHEQETMLVKT---AKNAYQGAGESQEEEGLDGYTA 137
             | | :|    |   |          |:| | ::|   :::  :|:| | :  | :  ||
Db     277 LVTFGHDGRGHTLTRRRAKRSPKTRQEEEYMPMETRTQSQDVSRGSG-SSDYNGSELKTA 335

Qy     138 VGPLLARRKRSTGASSHCQKTSLRVNFEDIGWDSWIIAPKEYDAYECKGGCFFPLADDVT 197
                         |:|  | |:|:|:||  |||||| | |  | | | |||   :
Db     336 -----------------CKKHELYVSFQDLGWQDWIIAPKGYAANYCDGECSFPLNAHMN 378

Qy     198 PTKHAIVQTLVHLKFPTKVGKACCVPTKLSPISILYKDDMGVPTLKYHYEGMSVAECGC 256
             | ||||||||||  |  | | || ||||: ||:|| ||     ||  |  | |  |||
Db     379 ATNHAIVQTLVHLMNPEYVPKPCCAPTKLNAISVLYFDDNSNVILK-KYRNMVVRACGC 436


RESULT 6
US-09-784-911-6
; Sequence 6, Application US/09784911
; Patent No. US20020072115A1
; GENERAL INFORMATION:
;  APPLICANT: Harrison, Leonard C.
;  APPLICANT:  Jiang, Fang-Xu
;  APPLICANT:  Stanley, Edouard Guy
;  APPLICANT:  Gonez, Leonel Jorge
;  TITLE OF INVENTION: Pancreatic islet cell growth factors
;  FILE REFERENCE: Davies Collison Cave
;  CURRENT APPLICATION NUMBER: US/09/784,911
;  CURRENT FILING DATE:  2001-09-17
;  NUMBER OF SEQ ID NOS: 30
;  SOFTWARE: PatentIn version 2.1
; SEQ ID NO 6
;   LENGTH: 427
;   TYPE: PRT
;   ORGANISM: mouse
;   FEATURE:
;   NAME/KEY: UNSURE
;   LOCATION: (186)
;   OTHER INFORMATION: Xaa at position 186 is Tyr or His
US-09-784-911-6

  Query Match             25.6%;  Score 350;  DB 10;  Length 427;
  Best Local Similarity   36.8%;  Pred. No. 1.1e-26;
  Matches   89;  Conservative   32;  Mismatches   77;  Indels   44;  Gaps    9;

Qy      29 WETLEVSSAVKRWVRADSTTNKNKLEVT----VQSHRESCDTLDISVPPGSKN----LPF 80
           ||| :|| || || |          :|||     ::|:     :  |:| || :     |
Db     217 WETFDVSPAVLRWTREKQPNYGLAIEVTHLHQTRTHQGQHVRISRSLPQGSGDWAQLRPL 276

Qy      81 FVVFSNDRSNGTKETRLDLLKEMIGH-----EQETMLVKTAKNAYQGAGESQEEEGLDGY 135
             | | :|    |   ||    :    |       ||  |:            |:|:
Db     277 LVTFGHD-GRGHTLTRRRAKRSPKHHPQRSTRQEQKLI------------SEED------ 317

Qy     136 TAVGPLLARRKRSTGASSHCQKTSLRVNFEDIGWDSWIIAPKEYDAYECKGGCFFPLADD 195
                 | |||:    : :|:: || |:| |:|: ||:||  | |: | | | |||||
Db     318 ------LTRRKK----NKNCRRHSLYVDFSDVGWNDWIVAPPGYQAFYCHGDCPFPLADH 367

Qy     196 VTPTKHAIVQTLVHLKFPTKVGKACCVPTKLSPISILYKDDMGVPTLKYHYEGMSVAECG 255
           :  | |||||||||:    : : |||||||:|| ||:|| |:     || :|: | |  ||
Db     368 LNSTNHAIVQTLVN-SVNSSIPKACCVPTELSAISMLYLDEYDKVVLK-NYQEMVVEGCG 425

Qy     256 CR 257
           ||
Db     426 CR 427


RESULT 7
US-09-945-182-26
; Sequence 26, Application US/09945182
; Patent No. US20020160494A1
;   GENERAL INFORMATION:
;        APPLICANT: Celeste, Anthony J.
;                   Wozney, John
;                   Rosen, Vicki A.
;                   Wolfman, Neil
;                   Thomsen, Gerald H.
;                   Melton, Douglas A.
;        TITLE OF INVENTION: TENDON-INDUCING COMPOSITIONS
;        NUMBER OF SEQUENCES: 35
;        CORRESPONDENCE ADDRESS:
;             ADDRESSEE: GENETICS INSTITUTE, INC.
;             STREET: 87 CambridgePark Drive
;             CITY: Cambridge
;             STATE: Massachusetts
;             COUNTRY: USA
;             ZIP: 02140
;        COMPUTER READABLE FORM:
;             MEDIUM TYPE: Floppy disk
;             COMPUTER: IBM PC compatible
;             OPERATING SYSTEM: PC-DOS/MS-DOS
;             SOFTWARE: PatentIn Release #1.0, Version #1.25
;        CURRENT APPLICATION DATA:
;             APPLICATION NUMBER: US/09/945,182
;             FILING DATE: 31-Aug-2001
;             CLASSIFICATION: <Unknown>
;        PRIOR APPLICATION DATA:
;             APPLICATION NUMBER: 08/808,324
;             FILING DATE: <Unknown>
;        ATTORNEY/AGENT INFORMATION:
;             NAME: Lazar, Steven R.
;             REGISTRATION NUMBER: 32,618
;             REFERENCE/DOCKET NUMBER: 5202-D
;        TELECOMMUNICATION INFORMATION:
;             TELEPHONE: 617 498-8260
;             TELEFAX: 617 876-5851
;   INFORMATION FOR SEQ ID NO: 26:
;        SEQUENCE CHARACTERISTICS:
;             LENGTH: 321 amino acids
;             TYPE: amino acid
;             TOPOLOGY: linear
;        MOLECULE TYPE: protein
;        SEQUENCE DESCRIPTION: SEQ ID NO: 26:
US-09-945-182-26

  Query Match             25.5%;  Score 349.5;  DB 9;  Length 321;
  Best Local Similarity   34.6%;  Pred. No. 8.3e-27;
  Matches  100;  Conservative   34;  Mismatches   94;  Indels   61;  Gaps   12;

Qy       2 VLEDSETWD----QATGTKTFLVSQDIRDEGWE--TLEVSSAVKRWVRADSTTNKNKLEV 55
```

```
          :| |: | |       | : | | | :| : |:   ||: :|   |   |:     : |
Db     61 LLLDARTLDPQGAPPAGWEVFDVWQGLRHQPWKQLCLELRAA---WGELDA----GEAEA 113

Qy     56 TVQSHRESCDTLDISVPPGSKNLPF------------FVVFSNDRSNGTKETRLDLLKEM 103
            :  ::         ||  ::| |             |||        |:  | :|  ||
Db    114 RARGPQQ-------PPPPDLRSLGFGRRVRPPQERALLVVF-------TRSQRKNLFAEM 159

Qy    104 IGHEQETMLVKTAKNAYQGAGESQEEEGLDGYTAVGPLL----ARRKRSTGASSH----- 154
             ||    : :|: |  |||          |     | |     ||:|:  || |
Db    160 --REQ----LGSAEAAGPGAGAEGSWPPPSGAPDARPWLPSPGRRRRRTAFASRHGKRHG 213

Qy    155 ------CQKTSLRVNFEDIGWDSWIIAPKEYDAYECKGGCFFPLADDVTPTKHAIVQTLV 208
                | |  | |||:::||| ||||| ||:|| |:| | |||   : || |||:|||:
Db    214 KKSRLRCSKKPLHVNFKELGWDDWIIAPLEYEAYHCEGVCDFPLRSHLEPTNHAIIQTLM 273

Qy    209 HLKFPTKVGKACCVPTKLSPISILYKDDMGVPTLKYHYEGMSVAECGCR 257
          :   |     :|||||||:||||||  | |   :   || | |  ||||
Db    274 NSMDPGSTPPSCCVPTKLTPISILY-IDAGNNVVYKQYEDMVVESCGCR 321


RESULT 8
US-09-804-625-6
; Sequence 6, Application US/09804625
; Publication No. US20030049826A1
;   GENERAL INFORMATION:
;        APPLICANT: Wang, Elizabeth A.
;                   Rosen, Vicki A.
;                   Wozney, John M
;        TITLE OF INVENTION: No. US20030049826A1el BMP Products
;        NUMBER OF SEQUENCES: 6
;        CORRESPONDENCE ADDRESS:
;             ADDRESSEE: LEGAL AFFAIRS, GENETICS INSTITUTE, INC.
;             STREET: 87 CAMBRIDGEPARK DRIVE
;             CITY: CAMBRIDGE
;             STATE: MA
;             COUNTRY: USA
;             ZIP: 02140
;        COMPUTER READABLE FORM:
;             MEDIUM TYPE: Floppy disk
;             COMPUTER: IBM PC compatible
;             OPERATING SYSTEM: PC-DOS/MS-DOS
;             SOFTWARE: PatentIn Release #1.0, Version #1.25
;        CURRENT APPLICATION DATA:
;             APPLICATION NUMBER: US/09/804,625
;             FILING DATE: 09-Mar-2001
;             CLASSIFICATION: <Unknown>
;        PRIOR APPLICATION DATA:
;             APPLICATION NUMBER: 08/925,779
;             FILING DATE: <Unknown>
;        ATTORNEY/AGENT INFORMATION:
;             NAME: Kapinos, Ellen J.
;             REGISTRATION NUMBER: 32,245
;             REFERENCE/DOCKET NUMBER: 5160C
;        TELECOMMUNICATION INFORMATION:
;             TELEPHONE: 617-876-1170
;             TELEFAX: 617-876-5851
;   INFORMATION FOR SEQ ID NO: 6:
;        SEQUENCE CHARACTERISTICS:
;             LENGTH: 408 amino acids
;             TYPE: amino acid
;             TOPOLOGY: linear
;        MOLECULE TYPE: protein
;        SEQUENCE DESCRIPTION: SEQ ID NO: 6:
US-09-804-625-6

  Query Match             25.2%;  Score 345;  DB 9;  Length 408;
  Best Local Similarity   35.9%;  Pred. No. 3.3e-26;
  Matches   85;  Conservative   28;  Mismatches   76;  Indels   48;  Gaps    7;

Qy     29 WETLEVSSAVKRWVRADSTTNKNKLEVT----VQSHRESCDTLDISVPPGSKN----LPF 80
          ||| :|| || || |         :|||     ::|:     :  |:| || |     |
Db    212 WETFDVSPAVLRWTREKQPNYGLAIEVTHLHQTRTHQGQHVRISRSLPQGSGNWAQLRPL 271

Qy     81 FVVFSNDRSNGTKETRLDLLKEMIGHEQETMLVKTAKNAYQGAGESQEEEGLDGYTAVGP 140
          | | :|   |   ||    |   | :
Db    272 LVTFGHD-GRGHALTRRRRAKRSPKHHSQR---------------------------- 300

Qy    141 LLARRKRSTGASSHCQKTSLRVNFEDIGWDSWIIAPKEYDAYECKGGCFFPLADDVTPTK 200
            ||:|     : :|:: || |:| |:||: ||:||  | |: | | | |||||  :  |
Db    301 --ARKK-----NKNCRRHSLYVDFSDVGWNDWIVAPPGYQAFYCHGDCPFPLADHLNSTN 353

Qy    201 HAIVQTLVHLKFPTKVGKACCVPTKLSPISILYKDDMGVPTLKYHYEGMSVAECGCR 257
          |||||||:    : : |||||||:|| ||:|| |:     || :|: | |  ||||
Db    354 HAIVQTLVN-SVNSSIPKACCVPTELSAISMLYLDEYDKVVLK-NYQEMVVEGCGCR 408


RESULT 9
US-09-749-728B-69
; Sequence 69, Application US/09749728B
; Patent No. US20020142457A1
; GENERAL INFORMATION:
;  APPLICANT: Umezawa, Akihiro
;  APPLICANT:  Hata, Jun-Ichi
;  APPLICANT:  Fukuda, Keiichi
;  APPLICANT:  Ogawa, Satoshi
;  APPLICANT:  Sakurada, Kazuhiro
;  APPLICANT:  Gojo, Satoshi
;  APPLICANT:  Yamada, Yoji
;  TITLE OF INVENTION: THE CELL HAVING THE POTENTIALITY OF DIFFERENTIATION INTO CARDI
;  FILE REFERENCE: 00766.000043
;  CURRENT APPLICATION NUMBER: US/09/749,728B
;  CURRENT FILING DATE:  2001-09-17
;  PRIOR APPLICATION NUMBER: H11-372826
;  PRIOR FILING DATE: 1999-12-28
;  PRIOR APPLICATION NUMBER: PCT-JP00-01148
;  PRIOR FILING DATE: 2000-02-28
;  PRIOR APPLICATION NUMBER: PCT-JP00-07741
;  PRIOR FILING DATE: 2000-11-02
;  NUMBER OF SEQ ID NOS: 80
;  SOFTWARE: PatentIn Ver.2.0
; SEQ ID NO 69
;   LENGTH: 408
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-749-728B-69

  Query Match             25.2%;  Score 345;  DB 10;  Length 408;
  Best Local Similarity   35.9%;  Pred. No. 3.3e-26;
  Matches   85;  Conservative   28;  Mismatches   76;  Indels   48;  Gaps    7;

Qy     29 WETLEVSSAVKRWVRADSTTNKNKLEVT----VQSHRESCDTLDISVPPGSKN----LPF 80
          ||| :|| || || |         :|||     ::|:     :  |:| || |     |
Db    212 WETFDVSPAVLRWTREKQPNYGLAIEVTHLHQTRTHQGQHVRISRSLPQGSGNWAQLRPL 271

Qy     81 FVVFSNDRSNGTKETRLDLLKEMIGHEQETMLVKTAKNAYQGAGESQEEEGLDGYTAVGP 140
          | | :|   |   ||    |   | :
Db    272 LVTFGHD-GRGHALTRRRRAKRSPKHHSQR---------------------------- 300

Qy    141 LLARRKRSTGASSHCQKTSLRVNFEDIGWDSWIIAPKEYDAYECKGGCFFPLADDVTPTK 200
            ||:|     : :|:: || |:| |:||: ||:||  | |: | | | |||||  :  |
Db    301 --ARKK-----NKNCRRHSLYVDFSDVGWNDWIVAPPGYQAFYCHGDCPFPLADHLNSTN 353

Qy    201 HAIVQTLVHLKFPTKVGKACCVPTKLSPISILYKDDMGVPTLKYHYEGMSVAECGCR 257
          |||||||:    : : |||||||:|| ||:|| |:     || :|: | |  ||||
Db    354 HAIVQTLVN-SVNSSIPKACCVPTELSAISMLYLDEYDKVVLK-NYQEMVVEGCGCR 408


RESULT 10
US-09-813-398-27
; Sequence 27, Application US/09813398
; Patent No. US20020169292A1
; GENERAL INFORMATION:
;  APPLICANT: Bruce D. Weintraub
;  APPLICANT:  Mariusz W. Szkudlinski
;  APPLICANT:  University of Maryland
;  TITLE OF INVENTION: CYSTINE KNOT GROWTH FACTOR MUTANTS
```

```
;  FILE REFERENCE: UOFMD.003C1
;  CURRENT APPLICATION NUMBER: US/09/813,398
;  CURRENT FILING DATE:  2001-03-20
;  PRIOR APPLICATION NUMBER: PCT/US99/05908
;  PRIOR FILING DATE: 1999-03-19
;  PRIOR APPLICATION NUMBER: PCT/US98/19772
;  PRIOR FILING DATE: 1998-09-22
;  NUMBER OF SEQ ID NOS: 41
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 27
;   LENGTH: 409
;   TYPE: PRT
;   ORGANISM: HOMO SAPIEN
US-09-813-398-27

  Query Match             25.2%;  Score 345;  DB 9;  Length 409;
  Best Local Similarity   35.9%;  Pred. No. 3.3e-26;
  Matches   85;  Conservative   28;  Mismatches   76;  Indels   48;  Gaps    7;

Qy      29 WETLEVSSAVKRWVRADSTTNKNKLEVT----VQSHRESCDTLDISVPPGSKN----LPF 80
           ||| :|| || || |          :|||      ::|:      :  |:| || '|      |
Db     213 WETFDVSPAVLRWTREKQPNYGLAIEVTHLHQTRTHQGQHVRISRSLPQGSGNWAQLRPL 272

Qy      81 FVVFSNDRSNGTKETRLDLLKEMIGHEQETMLVKTAKNAYQGAGESQEEEGLDGYTAVGP 140
           | | :|    |   ||    |    | :
Db     273 LVTFGHD-GRGHALTRRRRAKRSPKHHSQR---------------------------- 301

Qy     141 LLARRKRSTGASSHCQKTSLRVNFEDIGWDSWIIAPKEYDAYECKGGCFFPLADDVTPTK 200
            ||:|      : :|:: || |:| |:||: ||:||  | |: | | | |||||  :  |
Db     302 --ARKK-----NKNCRRHSLYVDFSDVGWNDWIVAPPGYQAFYCHGDCPFPLADHLNSTN 354

Qy     201 HAIVQTLVHLKFPTKVGKACCVPTKLSPISILYKDDMGVPTLKYHYEGMSVAECGCR 257
           ||||||||:    : : |||||||:|| ||:|| |:     || :|: | |  ||||
Db     355 HAIVQTLVN-SVNSSIPKACCVPTELSAISMLYLDEYDKVVLK-NYQEMVVEGCGCR 409


RESULT 11
US-09-945-182-32
; Sequence 32, Application US/09945182
; Patent No. US20020160494A1
;   GENERAL INFORMATION:
;        APPLICANT: Celeste, Anthony J.
;                   Wozney, John
;                   Rosen, Vicki A.
;                   Wolfman, Neil
;                   Thomsen, Gerald H.
;                   Melton, Douglas A.
;        TITLE OF INVENTION: TENDON-INDUCING COMPOSITIONS
;        NUMBER OF SEQUENCES: 35
;        CORRESPONDENCE ADDRESS:
;             ADDRESSEE: GENETICS INSTITUTE, INC.
;             STREET: 87 CambridgePark Drive
;             CITY: Cambridge
;             STATE: Massachusetts
;             COUNTRY: USA
;             ZIP: 02140
;        COMPUTER READABLE FORM:
;             MEDIUM TYPE: Floppy disk
;             COMPUTER: IBM PC compatible
;             OPERATING SYSTEM: PC-DOS/MS-DOS
;             SOFTWARE: PatentIn Release #1.0, Version #1.25
;        CURRENT APPLICATION DATA:
;             APPLICATION NUMBER: US/09/945,182
;             FILING DATE: 31-Aug-2001
;             CLASSIFICATION: <Unknown>
;        PRIOR APPLICATION DATA:
;             APPLICATION NUMBER: 08/808,324
;             FILING DATE: <Unknown>
;        ATTORNEY/AGENT INFORMATION:
;             NAME: Lazar, Steven R.
;             REGISTRATION NUMBER: 32,618
;             REFERENCE/DOCKET NUMBER: 5202-D
;        TELECOMMUNICATION INFORMATION:
;             TELEPHONE: 617 498-8260
;             TELEFAX: 617 876-5851
;   INFORMATION FOR SEQ ID NO: 32:
;        SEQUENCE CHARACTERISTICS:
;             LENGTH: 263 amino acids
;             TYPE: amino acid
;             TOPOLOGY: linear
;        MOLECULE TYPE: protein
;        SEQUENCE DESCRIPTION: SEQ ID NO: 32:
US-09-945-182-32

  Query Match             24.7%;  Score 337.5;  DB 9;  Length 263;
  Best Local Similarity   35.4%;  Pred. No. 9.8e-26;
  Matches   96;  Conservative   28;  Mismatches   86;  Indels   61;  Gaps   12;

Qy      16 KTFLVSQDIRDEGWE--TLEVSSAVKRWVRADSTTNKNKLEVTVQSHRESCDTLDISV-- 71
           |:| | | :| : |:   ||: :|   |   |      :     |       ||:
Db      25 KSFDVWQGLRPQPWKQLCLELRAA---WGELDXGDTGARARGPQQP-----PPLDLRSLG 76

Qy      72 ------PPGSKNLPFFVVFSNDRSNGTKETRLDLLKEMIGHEQETMLVKTAKNAYQGAGE 125
                 ||  : |   |||       |:  | :|  ||  |||             |: |
Db      77 FGRRVRPPQERAL--LVVF-------TRSQRKNLFTEM--HEQ-----------LGSAE 113

Qy     126 SQEEEGL----DGYTAVGPLL----ARRKRSTGASSH----------CQKTSLRVNFED 166
           :   ||       |   | |      ||:|:  || |          | :  | |||::
Db     114 AAGAEGSCPAPSGSPDTGSWLPSPGRRRRRTAFASRHGKRHGKKSRLRCSRKPLHVNFKE 173

Qy     167 IGWDSWIIAPKEYDAYECKGGCFFPLADDVTPTKHAIVQTLVHLKFPTKVGKACCVPTKL 226
           :||| ||||| ||:|| |:| | |||   : || |||:|||::   |     :|||||||
Db     174 LGWDDWIIAPLEYEAYHCEGVCDFPLRSHLEPTNHAIIQTLMNSMDPGSTPPSCCVPTKL 233

Qy     227 SPISILYKDDMGVPTLKYHYEGMSVAECGCR 257
           :|||||  | |   :   || | |  ||||
Db     234 TPISILY-IDAGNNXVYKQYEDMVVESCGCR 263


RESULT 12
US-09-945-182-28
; Sequence 28, Application US/09945182
; Patent No. US20020160494A1
;   GENERAL INFORMATION:
;        APPLICANT: Celeste, Anthony J.
;                   Wozney, John
;                   Rosen, Vicki A.
;                   Wolfman, Neil
;                   Thomsen, Gerald H.
;                   Melton, Douglas A.
;        TITLE OF INVENTION: TENDON-INDUCING COMPOSITIONS
;        NUMBER OF SEQUENCES: 35
;        CORRESPONDENCE ADDRESS:
;             ADDRESSEE: GENETICS INSTITUTE, INC.
;             STREET: 87 CambridgePark Drive
;             CITY: Cambridge
;             STATE: Massachusetts
;             COUNTRY: USA
;             ZIP: 02140
;        COMPUTER READABLE FORM:
;             MEDIUM TYPE: Floppy disk
;             COMPUTER: IBM PC compatible
;             OPERATING SYSTEM: PC-DOS/MS-DOS
;             SOFTWARE: PatentIn Release #1.0, Version #1.25
;        CURRENT APPLICATION DATA:
;             APPLICATION NUMBER: US/09/945,182
;             FILING DATE: 31-Aug-2001
;             CLASSIFICATION: <Unknown>
;        PRIOR APPLICATION DATA:
;             APPLICATION NUMBER: 08/808,324
;             FILING DATE: <Unknown>
;        ATTORNEY/AGENT INFORMATION:
;             NAME: Lazar, Steven R.
;             REGISTRATION NUMBER: 32,618
;             REFERENCE/DOCKET NUMBER: 5202-D
;        TELECOMMUNICATION INFORMATION:
```

```
;             TELEPHONE: 617 498-8260
;             TELEFAX: 617 876-5851
;    INFORMATION FOR SEQ ID NO: 28:
;         SEQUENCE CHARACTERISTICS:
;             LENGTH: 411 amino acids
;             TYPE: amino acid
;             TOPOLOGY: linear
;         MOLECULE TYPE: protein
;         SEQUENCE DESCRIPTION: SEQ ID NO: 28:
US-09-945-182-28

  Query Match             24.5%;  Score 334.5;  DB 9;  Length 411;
  Best Local Similarity   32.0%;  Pred. No. 3.7e-25;
  Matches   83;  Conservative   32;  Mismatches   85;  Indels   59;  Gaps    8;

Qy      17 TFLVSQDIRDEGWETLEVSSAVKRWVRADSTTNKNKLEVTVQSHRESCDTLDISVPPGSK 76
           | ||:|:     ||: :|: || ||       :   :||     ::       : :   |:
Db     194 TRLVNQNA--SRWESFDVTPAVMRWTAQGHANHGFVVEVAHLEEKQGVSKRHVRI---SR 248

Qy      77 NL-----------PFFVVFSNDRSNGTKETRLDLLKEMIGH-----EQETML--VKTAKN 118
           :|           |  | | :|                ||     |: | |   :||:
Db     249 SLHQDEHSWSQIRPLLVTFGHDGK--------------GHPLHKREKRTALAGTRTAQG 293

Qy     119 AYQGAGESQEEEGLDGYTAVGPLLARRKRSTGASSHCQKTSLRVNFEDIGWDSWIIAPKE 178
           :  |||      |                    | | :  | |:|:::||| ||||| :
Db     294 SGGGAGRGHGRRG------------------RSRCSRKPLHVDFKELGWDDWIIAPLD 333

Qy     179 YDAYECKGGCFFPLADDVTPTKHAIVQTLVHLKFPTKVGKACCVPTKLSPISILYKDDMG 238
           |:|| |:| | |||   : || |||:|||::   |     :|||| :||||||| |
Db     334 YEAYHCEGLCDFPLRSHLEPTNHAIIQTLLNSMAPDAAPASCCVPARLSPISILYIDAAN 393

Qy     239 VPTLKYHYEGMSVAECGCR 257
             |  || | |  ||||
Db     394 NVVYK-QYEDMVVEACGCR 411


RESULT 13
US-09-730-772-14
; Sequence 14, Application US/09730772
; Patent No. US20010011131A1
;  GENERAL INFORMATION:
;    APPLICANT:  Luyten, Frank P.
;    APPLICANT:  Moos, Jr., Malcolm
;    APPLICANT:  Chang, Steven Chao-Huan
;    TITLE OF INVENTION:  CARTILAGE-DERIVED MORPHOGENETIC
;    TITLE OF INVENTION:  PROTEINS
;    NUMBER OF SEQUENCES:  24
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Knobbe, Martens, Olson & Bear
;      STREET:  620 Newport Center Drive, 16th Floor
;      CITY:  Newport Beach
;      STATE:  CA
;      COUNTRY:  U.S.A.
;      ZIP:  92660
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  Windows
;      SOFTWARE:  FastSEQ for Windows Version 2.0b
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/730,772
;      FILING DATE:
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/836,081
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Bartfeld, Neil S
;      REGISTRATION NUMBER:  39,901
;      REFERENCE/DOCKET NUMBER:  NIH099.001APC
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  619-235-8550
;      TELEFAX:  619-235-0176
;      TELEX:
;  INFORMATION FOR SEQ ID NO:  14:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  436 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
US-09-730-772-14

  Query Match             24.1%;  Score 330;  DB 10;  Length 436;
  Best Local Similarity   36.7%;  Pred. No. 1.1e-24;
  Matches   81;  Conservative   23;  Mismatches   67;  Indels   50;  Gaps    7;

Qy      72 PPGSKNLPF-----------FVVFSNDRSNGTKETRLDLLKEMIGHEQETMLVKTAKNA 119
           ||  ::| |             ||||  :    ::|    ::| :|   | :
Db     231 PPDLRSLGFGRRVRTPQERALLVVFSRSQ----RKTLFAEMREQLGSATEVV-------- 278

Qy     120 YQGAGESQEEEG---------LDGYTAVG---PLLARRKRSTGASSH-----------CQ 156
           | |   | |             |   |   |   ||:|:  || |           |
Db     279 --GPGGGAEGSGPPPPPPPPPPSGTPDAGLWSPSPGRRRRTAFASRHGKRHGKKSRLRCS 336

Qy     157 KTSLRVNFEDIGWDSWIIAPKEYDAYECKGGCFFPLADDVTPTKHAIVQTLVHLKFPTKV 216
           |  | |||:::|||  ||||| ||:|| |:| | |||   : || |||:|||::   |
Db     337 KKPLHVNFKELGWDDWIIAPLEYEAYHCEGVCDFPLRSHLEPTNHAIIQTLMNSMDPGST 396

Qy     217 GKACCVPTKLSPISILYKDDMGVPTLKYHYEGMSVAECGCR 257
             :|||||||:||||||  | |   :   || | |  ||||
Db     397 PPSCCVPTKLTPISILY-IDAGNNVVYNEYEEMVVESCGCR 436


RESULT 14
US-09-735-849-14
; Sequence 14, Application US/09735849
; Patent No. US20010037017A1
;  GENERAL INFORMATION:
;    APPLICANT:  Luyten, Frank P.
;    APPLICANT:  Moos, Jr., Malcolm
;    APPLICANT:  Chang, Steven Chao-Huan
;    TITLE OF INVENTION:  CARTILAGE-DERIVED MORPHOGENETIC
;    TITLE OF INVENTION:  PROTEINS
;    NUMBER OF SEQUENCES:  24
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Knobbe, Martens, Olson & Bear
;      STREET:  620 Newport Center Drive, 16th Floor
;      CITY:  Newport Beach
;      STATE:  CA
;      COUNTRY:  U.S.A.
;      ZIP:  92660
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  Windows
;      SOFTWARE:  FastSEQ for Windows Version 2.0b
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/735,849
;      FILING DATE:
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/836,081
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Bartfeld, Neil S
;      REGISTRATION NUMBER:  39,901
;      REFERENCE/DOCKET NUMBER:  NIH099.001APC
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  619-235-8550
;      TELEFAX:  619-235-0176
;      TELEX:
;  INFORMATION FOR SEQ ID NO:  14:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  436 amino acids
;      TYPE:  amino acid
```

```
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
US-09-735-849-14

  Query Match             24.1%;  Score 330;  DB 10;  Length 436;
  Best Local Similarity   36.7%;  Pred. No. 1.1e-24;
  Matches   81;  Conservative   23;  Mismatches   67;  Indels   50;  Gaps    7;

Qy      72 PPGSKNLPF-----------FVVFSNDRSNGTKETRLDLLKEMIGHEQETMLVKTAKNA 119
           ||  ::| |             ||||  :    ::|    ::| :|   | :
Db     231 PPDLRSLGFGRRVRTPQERALLVVFSRSQ----RKTLFAEMREQLGSATEVV-------- 278

Qy     120 YQGAGESQEEEG---------LDGYTAVG---PLLARRKRSTGASSH----------CQ 156
             | |   |  |           |    |   |   ||:|:  || |           |
Db     279 --GPGGGAEGSGPPPPPPPPPPSGTPDAGLWSPSPGRRRRTAFASRHGKRHGKKSRLRCS 336

Qy     157 KTSLRVNFEDIGWDSWIIAPKEYDAYECKGGCFFPLADDVTPTKHAIVQTLVHLKFPTKV 216
           |  | |||:::||| ||||| ||:|| |:| | |||   : || |||:|||::   |
Db     337 KKPLHVNFKELGWDDWIIAPLEYEAYHCEGVCDFPLRSHLEPTNHAIIQTLMNSMDPGST 396

Qy     217 GKACCVPTKLSPISILYKDDMGVPTLKYHYEGMSVAECGCR 257
             :|||||||:||||||  | |   :   || | |  ||||
Db     397 PPSCCVPTKLTPISILY-IDAGNNVVYNEYEEMVVESCGCR 436


RESULT 15
US-09-784-911-10
; Sequence 10, Application US/09784911
; Patent No. US20020072115A1
; GENERAL INFORMATION:
;  APPLICANT: Harrison, Leonard C.
;  APPLICANT:  Jiang, Fang-Xu
;  APPLICANT:  Stanley, Edouard Guy
;  APPLICANT:  Gonez, Leonel Jorge
;  TITLE OF INVENTION: Pancreatic islet cell growth factors
;  FILE REFERENCE: Davies Collison Cave
;  CURRENT APPLICATION NUMBER: US/09/784,911
;  CURRENT FILING DATE:  2001-09-17
;  NUMBER OF SEQ ID NOS: 30
;  SOFTWARE: PatentIn version 2.1
; SEQ ID NO 10
;   LENGTH: 451
;   TYPE: PRT
;   ORGANISM: mouse
US-09-784-911-10

  Query Match             23.5%;  Score 322;  DB 10;  Length 451;
  Best Local Similarity   32.9%;  Pred. No. 7.3e-24;
  Matches   93;  Conservative   35;  Mismatches   91;  Indels   64;  Gaps   12;

Qy       5 DSETW--------DQATGTKTFLVSQDIR-----DEGWETLEVSSAVKRWVRADSTTNKN 51
           |:||:         : :| :: |   | |      :|||   ::::    ||       ::
Db     185 DNETFQITVYQVLQEHSGRESDLFLLDSRTIWASEEGWLVFDITATSNHWV----VNPRH 240

Qy      52 KLEVTVQSHRESCDTLD-ISVPP---------GSKN-LPFFVVFSNDRSNGTKETRLDLL 100
            | : :     | :|||  |: |          | :|  || | |          |  |  :
Db     241 NLGLQL-----SVETLDGQSINPKLAGLIGRHGPQNKQPFMVAFFK-----ATEVHLRSI 290

Qy     101 KEMIGHEQETMLVK-------TAKNAYQGAGESQEEEGLDGYTAVGPLLARRKRSTGASS 153
           :   |  ||  |:         :::  :|:| | :  | :  ||
Db     291 RSTGGTRQEQKLISEEDLTRTQSQDVSRGSG-SSDYNGSELKTA---------------- 333

Qy     154 HCQKTSLRVNFEDIGWDSWIIAPKEYDAYECKGGCFFPLADDVTPTKHAIVQTLVHLKFP 213
            |:|  | |:|:|:||  |||||| | |  | | | |||   :  | ||||||||||  |
Db     334 -CKKHELYVSFQDLGWQDWIIAPKGYAANYCDGECSFPLNAHMNATNHAIVQTLVHLMNP 392

Qy     214 TKVGKACCVPTKLSPISILYKDDMGVPTLKYHYEGMSVAECGC 256
             | | || ||||: ||:|| ||     ||  |  | |  |||
Db     393 EYVPKPCCAPTKLNAISVLYFDDNSNVILK-KYRNMVVRACGC 434


Search completed: April  2, 2003, 15:00:43
```

Job time : 37 secs

OM protein - protein search, using sw model

Run on: April 2, 2003, 14:40:05 ; Search time 75 Seconds
(without alignments)
456.605 Million cell updates/sec

Title: US-10-002-278-9
Perfect score: 1368
Sequence: 1 DVLEDSETWDQATGTKTFLV.....GVPTLKYHYEGMSVAECGCR 257

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_101002:*
1: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1980.DAT:*
2: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1981.DAT:*
3: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1982.DAT:*
4: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1983.DAT:*
5: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1984.DAT:*
6: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1985.DAT:*
7: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1986.DAT:*
8: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1987.DAT:*
9: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1988.DAT:*
10: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1989.DAT:*
11: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1990.DAT:*
12: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1991.DAT:*
13: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1992.DAT:*
14: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1993.DAT:*
15: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1994.DAT:*
16: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1995.DAT:*
17: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1996.DAT:*
18: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1997.DAT:*
19: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1998.DAT:*
20: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1999.DAT:*
21: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA2000.DAT:*
22: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA2001.DAT:*
23: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	1336.5	97.7	428	14	AAR31372	Murine bone morpho
2	1336.5	97.7	428	17	AAR86905	Murine BMP-9. Mus
3	1336.5	97.7	428	21	AAY82271	Murine bone morpho
4	1146.5	83.8	429	22	AAE12410	Albumin fusion pro
5	1146.5	83.8	429	22	AAU14195	Human novel protei
6	1146.5	83.8	429	22	AAU14431	Human novel protei
7	808.5	59.1	427	16	AAR68607	Dorsalin I. Gallu
8	634	46.3	150	17	AAR86903	Human BMP-9 polype
9	634	46.3	150	21	AAY82272	Human bone morphog
10	619	45.2	150	14	AAR31374	Human bone morphog
11	596	43.6	110	23	AAM51930	Human TGFbeta prot
12	519.5	38.0	312	16	AAR66200	Part of bovine bon
13	516.5	37.8	424	16	AAR66202	Complete human bon
14	516.5	37.8	424	21	AAY92029	Human bone morphog
15	508	37.1	421	20	AAY06304	Mouse pancreatic-d
16	469	34.3	103	20	AAY16715	WO9914235 Seq ID N
17	469	34.3	103	21	AAB09533	Chicken Dorsalin,
18	469	34.3	103	21	AAB02799	Chicken DORSALIN a
19	469	34.3	103	21	AAY92568	Dorsalin finger-1-
20	400.5	29.3	393	20	AAY06302	Mouse pancreatic-d
21	384.5	28.1	108	23	AAM51926	Human TGFbeta prot
22	361	26.4	419	22	AAG66568	BMP 4/4 glu polype
23	354	25.9	437	22	AAG66569	BMP 4/6 glu polype
24	350	25.6	427	22	AAG66570	BMP 4/6 myc polype
25	349.5	25.5	321	16	AAR78730	Human mature VL-1
26	349.5	25.5	321	18	AAW26591	Human bone morphog
27	349.5	25.5	321	22	AAE10982	Human full length
28	349.5	25.5	455	22	AAM50216	Human growth/diffe
29	349.5	25.5	455	23	AAU79173	Human growth/diffe
30	349.5	25.5	455	23	AAE17604	Human extracellula
31	348.5	25.5	417	16	AAR79476	Mouse BMP-4. Mus
32	345	25.2	408	9	AAP80620	Human Bone Morphog
33	345	25.2	408	12	AAR14242	Human BMP-2B in la
34	345	25.2	408	13	AAR29282	Human pre-pro-BMP-
35	345	25.2	408	14	AAR44748	Osteogenic protein
36	345	25.2	408	14	AAR31373	Human bone morphog
37	345	25.2	408	14	AAR36733	Human BMP-4. Homo
38	345	25.2	408	15	AAR51654	Prepro human CBMP2
39	345	25.2	408	17	AAR85761	Human CBMP2B. Hom
40	345	25.2	408	17	AAR86906	Human BMP-4. Homo
41	345	25.2	408	18	AAW24850	Human bone morphog
42	345	25.2	408	18	AAW15405	Human BMP-2B. Hom
43	345	25.2	408	19	AAW44304	Human osteogenic p
44	345	25.2	408	20	AAY43112	Human prepro CBMP2
45	345	25.2	408	20	AAW89681	Human osteogenic p

ALIGNMENTS

RESULT 1
AAR31372
ID AAR31372 standard; Protein; 428 AA.
XX
AC AAR31372;
XX
DT 24-JUN-1993 (first entry)
XX
DE Murine bone morphogenic protein-9 (BMP-9).
XX
KW bone formation; cartilage formation; wound healing; tissue repair;
KW surgery; fracture treatment; periodontal disease; osteoporosis;
KW increase neuronal survival; transplantation; nerves; nervous system.
XX
OS Mus musculus.
XX
FH Key Location/Qualifiers
FT Protein 319..427
FT /label= mature peptide
XX
PN WO9300432-A.
XX
PD 07-JAN-1993.
XX
PF 25-JUN-1992; 92WO-US05374.
XX
PR 25-JUN-1991; 91US-0720590.
XX
PA (GEMY) GENETICS INST INC.
XX
PI Celeste AJ, Wozney JM;
XX
DR WPI; 1993-036379/04.

```
DR   N-PSDB; AAQ35243.
XX
PT   New BMP-9 polypeptide(s) induce bone and cartilage formation -
PT   used for treating osteoporosis and fractures, healing wounds and
PT   increasing neuronal survival
XX
PS   Claim 1; Fig 1; 60pp; English.
XX
CC   This sequence is thought to be the primary translation product of
CC   murine bone morphogenic protein. Based on the knowledge of other
CC   BMP's and proteins within the TGF-Beta family it is predicted that
CC   the precursor polypeptide would be cleaved at the multibasic
CC   sequence Arg-Arg-Lys-Arg in agreement with the proposed concensus
CC   proteolytic processing sequence of Arg-X-X-Arg to generate a 110
CC   amino acid mature peptide. It is expected that processing into the
CC   mature form will involve dimerisation and removal of the N-terminal
CC   region in a manner analogous to the related protein TGF-beta,
CC   giving a homodimer of 2 subunits each with a molecular weight of
CC   approx. 12,000 Da. The protein can be used to induce bone and/or
CC   cartilage formation and in wound healing and tissue repair. It
CC   can be used in surgery of the treatment of fractures, periodontal
CC   disease or osteoporosis. The protein can also increase neuronal
CC   survival and can be used in transplantation and treatment of
CC   conditions exhibiting a decrease in neuronal survival.
XX
SQ   Sequence   428 AA;

  Query Match             97.7%;  Score 1336.5;  DB 14;  Length 428;
  Best Local Similarity   98.4%;  Pred. No. 9.9e-129;
  Matches  253;  Conservative    2;  Mismatches    1;  Indels    1;  Gaps    1;

Qy       1 DVLEDSETWDQATGTKTFLVSQDIRDEGWETLEVSSAVKRWVRADSTTNKNKLEVTVQSH 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     173 DVLEDSETWDQATGTKTFLVSQDIRDEGWETLEVSSAVKRWVRADSTTNKNKLEVTVQSH 232

Qy      61 RESCDTLDISVPPGSKNLPFFVVFSNDRSNGTKETRLDLLKEMIGHEQETMLVKTAKNAY 120
           |||||||||||||||||||||||||||||||||||: ||||||||||||||||||||||
Db     233 RESCDTLDISVPPGSKNLPFFVVFSNDRSNGTKETRLE-LKEMIGHEQETMLVKTAKNAY 291

Qy     121 QGAGESQEEEGLDGYTAVGPLLARRKRSTGASSHCQKTSLRVNFEDIGWDSWIIAPKEYD 180
           | ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     292 QVAGESQEEEGLDGYTAVGPLLARRKRSTGASSHCQKTSLRVNFEDIGWDSWIIAPKEYD 351

Qy     181 AYECKGGCFFPLADDVTPTKHAIVQTLVHLKFPTKVGKACCVPTKLSPISILYKDDMGVP 240
           |||||||||||||||||||||||||||||:||||||||||||||||||||||||||||||
Db     352 AYECKGGCFFPLADDVTPTKHAIVQTLVHLEFPTKVGKACCVPTKLSPISILYKDDMGVP 411

Qy     241 TLKYHYEGMSVAECGCR 257
           |||||||||||||||||
Db     412 TLKYHYEGMSVAECGCR 428


RESULT 2
AAR86905
ID   AAR86905 standard; Protein; 428 AA.
XX
AC   AAR86905;
XX
DT   10-MAY-1996  (first entry)
XX
DE   Murine BMP-9.
XX
KW   BMP-9; bone; cartilage; wound healing; liver; tissue repair.
XX
OS   Mus musculus strain C57B46xCBA.
XX
FH   Key             Location/Qualifiers
FT   Cleavage-site   318..319
FT                   /note= "proteolytic cleavage site"
FT   Protein         319..428
FT                   /label= Mat_protein
XX
PN   WO9533830-A1.
XX
PD   14-DEC-1995.
XX
PF   05-JUN-1995;   95WO-US07084.
XX
PR   06-JUN-1994;   94US-0254353.
XX
PA   (GEMY ) GENETICS INST INC.
XX
PI   Celeste AJ,  Rosen VA,  Song JJ,  Thies SR,  Wozney JM;
XX
DR   WPI; 1996-040235/04.
DR   N-PSDB; AAQ74084.
XX
PT   BMP-9 protein able to induce bone and cartilage formation - also
PT   useful in wound healing, tissue repair and for inducing hepatic
PT   growth and function
XX
PS   Example 1; Page 36-38; 75pp; English.
XX
CC   Murine BMP-9 protein (AAR86905) is the product of a cDNA clone,
CC   ML14a (AAQ74084), isolated from a mouse liver cDNA library.  The
CC   primary translation product is cleaved to yield a 110-amino acid
CC   protein that forms a mature, homodimeric, active species.
CC   Recombinant BMP-9 is obtd. by expression of the cDNA in procaryotic
CC   or eucaryotic hosts.  It has therapeutic applns. in the
CC   formation of bone and cartilage, wound healing and tissue repair.
XX
SQ   Sequence   428 AA;

  Query Match             97.7%;  Score 1336.5;  DB 17;  Length 428;
  Best Local Similarity   98.4%;  Pred. No. 9.9e-129;
  Matches  253;  Conservative    2;  Mismatches    1;  Indels    1;  Gaps    1;

Qy       1 DVLEDSETWDQATGTKTFLVSQDIRDEGWETLEVSSAVKRWVRADSTTNKNKLEVTVQSH 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     173 DVLEDSETWDQATGTKTFLVSQDIRDEGWETLEVSSAVKRWVRADSTTNKNKLEVTVQSH 232

Qy      61 RESCDTLDISVPPGSKNLPFFVVFSNDRSNGTKETRLDLLKEMIGHEQETMLVKTAKNAY 120
           |||||||||||||||||||||||||||||||||||: ||||||||||||||||||||||
Db     233 RESCDTLDISVPPGSKNLPFFVVFSNDRSNGTKETRLE-LKEMIGHEQETMLVKTAKNAY 291

Qy     121 QGAGESQEEEGLDGYTAVGPLLARRKRSTGASSHCQKTSLRVNFEDIGWDSWIIAPKEYD 180
           | ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     292 QVAGESQEEEGLDGYTAVGPLLARRKRSTGASSHCQKTSLRVNFEDIGWDSWIIAPKEYD 351

Qy     181 AYECKGGCFFPLADDVTPTKHAIVQTLVHLKFPTKVGKACCVPTKLSPISILYKDDMGVP 240
           |||||||||||||||||||||||||||||:||||||||||||||||||||||||||||||
Db     352 AYECKGGCFFPLADDVTPTKHAIVQTLVHLEFPTKVGKACCVPTKLSPISILYKDDMGVP 411

Qy     241 TLKYHYEGMSVAECGCR 257
           |||||||||||||||||
Db     412 TLKYHYEGMSVAECGCR 428


RESULT 3
AAY82271
ID   AAY82271 standard; Protein; 428 AA.
XX
AC   AAY82271;
XX
DT   16-JUN-2000  (first entry)
XX
DE   Murine bone morphogenetic protein 9 SEQ ID NO:2.
XX
KW   Bone morphogenetic protein 9; BMP-9; bone defect; cartilage defect;
KW   wound healing; tissue repair; hepatic growth; antiinfertility;
KW   osteopathic; antiarthritic; vulnerary.
XX
OS   Mus musculus.
XX
FH   Key             Location/Qualifiers
FT   Peptide         1..318
```

```
FT                   /label= Signal
FT   Protein         319..428
FT                   /label= Bone_morphogenetic_protein-9
XX
PN   US6034062-A.
XX
PD   07-MAR-2000.
XX
PF   13-MAR-1997;   97US-0815652.
XX
PR   13-MAR-1997;   97US-0815652.
XX
PA   (GEMY ) GENETICS INST INC.
XX
PI   Thies RS,  Song JJ;
XX
DR   WPI; 2000-255707/22.
DR   N-PSDB; AAZ95746.
XX
PT   Novel bone-morphogenetic protein (BMP)-9 polypeptides, useful in the
PT   treatment of bone and cartilage defects and in wound healing and tissue
PT   repair, as well as hepatic growth and function  -
XX
PS   Claim 1; Column 33-36; 36pp; English.
XX
CC   A method has been developed for decreasing proliferation of kidney or
CC   lung epithelial cells, or for inhibiting kidney tubule formation, in a
CC   patient. The method comprises administering an amount of a bone
CC   morphogenetic protein 9 (BMP-9) protein. BMP-9 polypeptides from the
CC   invention may be used to induce bone and cartilage formation or other
CC   connective tissue formation, in periodontal disease, in wound healing
CC   and tissue repair, hepatic growth and function, and in regulating lung
CC   and kidney cell function. The polypeptides can be used for healing of
CC   bone fractures, open fracture reduction, improved fixation of artificial
CC   joints, tooth repair processes, treatment of osteoporosis and
CC   osteoarthritis, and as a dietary supplement or as a component of cell
CC   culture media. BMP-9 may also be used for repair and regenerating of
CC   liver cells. The BMP-9 polypeptides may also have angiogenic,
CC   chemotactic and/or chemoattractant properties, and may induce collagen
CC   synthesis, fibrosis, differentiation processes, cell proliferative
CC   responses, cell adhesion responses, and migration. When dimerized, the
CC   BMP-9 polypeptide is expected to demonstrate effects on the production
CC   of follicle stimulating hormone (FSH), and so may be used as a
CC   contraceptive, as a fertility inducing therapeutic, and for advancing
CC   the onset of fertility in sexually immature mammals, so as to increase
CC   the lifetime reproductive performance of domestic animals. The present
CC   sequence represents murine BMP-9.
XX
SQ   Sequence   428 AA;

  Query Match             97.7%;  Score 1336.5;  DB 21;  Length 428;
  Best Local Similarity   98.4%;  Pred. No. 9.9e-129;
  Matches  253;  Conservative    2;  Mismatches    1;  Indels    1;  Gaps    1;

Qy       1 DVLEDSETWDQATGTKTFLVSQDIRDEGWETLEVSSAVKRWVRADSTTNKNKLEVTVQSH 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     173 DVLEDSETWDQATGTKTFLVSQDIRDEGWETLEVSSAVKRWVRADSTTNKNKLEVTVQSH 232

Qy      61 RESCDTLDISVPPGSKNLPFFVVFSNDRSNGTKETRLDLLKEMIGHEQETMLVKTAKNAY 120
          ||||||||||||||||||||||||||||||||||: ||||||||||||||||||||||
Db     233 RESCDTLDISVPPGSKNLPFFVVFSNDRSNGTKETRLE-LKEMIGHEQETMLVKTAKNAY 291

Qy     121 QGAGESQEEEGLDGYTAVGPLLARRKRSTGASSHCQKTSLRVNFEDIGWDSWIIAPKEYD 180
          | ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     292 QVAGESQEEEGLDGYTAVGPLLARRKRSTGASSHCQKTSLRVNFEDIGWDSWIIAPKEYD 351

Qy     181 AYECKGGCFFPLADDVTPTKHAIVQTLVHLKFPTKVGKACCVPTKLSPISILYKDDMGVP 240
          ||||||||||||||||||||||||||:||||||||||||||||||||||||||||||||||
Db     352 AYECKGGCFFPLADDVTPTKHAIVQTLVHLEFPTKVGKACCVPTKLSPISILYKDDMGVP 411

Qy     241 TLKYHYEGMSVAECGCR 257
          |||||||||||||||||
Db     412 TLKYHYEGMSVAECGCR 428
```

```
RESULT 4
AAE12410
ID   AAE12410 standard; Protein; 429 AA.
XX
AC   AAE12410;
XX
DT   18-DEC-2001  (first entry)
XX
DE   Albumin fusion protein related human protein #1.
XX
KW   Human; albumin; HA; immune system disorder; transplant rejection;
KW   blood related disorder; myocardial infarction; glomerulonephritis;
KW   hyperproliferative disorder; childhood acute myeloid leukaemia;
KW   renal cell carcinoma; cardiovascular disorder; vulnerary; melanoma;
KW   arrhythmia; respiratory disorder; non-allergic rhinitis; antileukaemic;
KW   neurological disease; Alzheimer's disease; endocrine disorder; measles;
KW   pheocytochroma; reproductive system disorder; neuroprotective; syphilis;
KW   infectious disease; gastrointestinal disorder; wound healing; nootropic;
KW   irritable bowel syndrome; HIV; human immunodeficiency virus infection;
KW   cytostatic; antiinflammatory; gene therapy; immunosuppressive; cardiant;
KW   antiarthritic; antirheumatic; renal disorder; antimicrobial.
XX
OS   Homo sapiens.
XX
PN   WO200179480-A1.
XX
PD   25-OCT-2001.
XX
PF   12-APR-2001; 2001WO-US11991.
XX
PR   12-APR-2000; 2000US-229358P.
PR   25-APR-2000; 2000US-199384P.
PR   21-DEC-2000; 2000US-256931P.
XX
PA   (HUMA-) HUMAN GENOME SCI INC.
XX
PI   Rosen CA,  Haseltine WA;
XX
DR   WPI; 2001-616756/71.
XX
PT   Albumin fusion proteins comprising a therapeutic protein and albumin,
PT   useful in the treating metastatic renal cell carcinoma, metastatic
PT   melanoma, malignant melanoma, renal cell carcinoma, HIV (human
PT   immunodeficiency virus) or infection -
XX
PS   Disclosure; Page 385-386; 394pp; English.
XX
CC   The invention relates to human albumin (HA) fusion proteins and their
CC   corresponding nucleic acid sequences. Therapeutic proteins fused to
CC   albumin or its fragments have an extended shelf-life. The albumin
CC   fusion proteins are useful in the treatment, prevention, diagnosis,
CC   and/or detection of diseases, disorders such as immune system
CC   disorders (e.g. transplant rejection), blood related disorders (e.g.
CC   myocardial infarction), hyperproliferative disorders (e.g. childhood
CC   acute myeloid leukaemia, metastatic renal cell carcinoma, metastatic
CC   melanoma, malignant melanoma, renal cell carcinoma), renal disorders
CC   (e.g. glomerulonephritis), cardiovascular disorders (e.g. arrhythmias),
CC   respiratory disorders (e.g. non-allergic rhinitis), neurological
CC   diseases (e.g. Alzheimer's disease), endocrine disorders (e.g.
CC   pheocytochroma), reproductive system disorders (e.g. syphilis),
CC   infectious diseases (e.g. measles), gastrointestinal disorders (e.g.
CC   irritable bowel syndrome), HIV (human immunodeficiency virus) infection
CC   and wound healing. Nucleic acids encoding albumin fusion protein is
CC   used in gene therapy. The present sequence is a human protein related
CC   to albumin fusion proteins.
XX
SQ   Sequence   429 AA;

  Query Match             83.8%;  Score 1146.5;  DB 22;  Length 429;
  Best Local Similarity   82.9%;  Pred. No. 3.6e-109;
  Matches  213;  Conservative   19;  Mismatches   24;  Indels    1;  Gaps    1;
```

```
Qy       1 DVLEDSETWDQATGTKTFLVSQDIRDEGWETLEVSSAVKRWVRADSTTNKNKLEVTVQSH 60
           |||: :: || || ||||||||||:||||||||||||||||||:||| :|||||||||:||
Db     174 DVLDGTDAWDSATETKTFLVSQDIQDEGWETLEVSSAVKRWVRSDSTKSKNKLEVTVESH 233

Qy      61 RESCDTLDISVPPGSKNLPFFVVFSNDRSNGTKETRLDLLKEMIGHEQETMLVKTAKNAY 120
           |: |||||||||||:||||||||||| |:|||||||: |:||| ||||::| | :|:
Db     234 RKGCDTLDISVPPGSRNLPFFVVFSNDHSSGTKETRLE-LREMISHEQESVLKKLSKDGS 292

Qy     121 QGAGESQEEEGLDGYTAVGPLLARRKRSTGASSHCQKTSLRVNFEDIGWDSWIIAPKEYD 180
             |||| || ||: | |  ||||||| || ||||||||||||||||||||||||||||:
Db     293 TEAGESSHEEDTDGHVAAGSTLARRKRSAGAGSHCQKTSLRVNFEDIGWDSWIIAPKEYE 352

Qy     181 AYECKGGCFPPLADDVTPTKHAIVQTLVHLKFPTKVGKACCVPTKLSPISILYKDDMGVP 240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||:|||||||||
Db     353 AYECKGGCFPPLADDVTPTKHAIVQTLVHLKFPTKVGKACCVPTKLSPISVLYKDDMGVP 412

Qy     241 TLKYHYEGMSVAECGCR 257
           |||||||||||||||||
Db     413 TLKYHYEGMSVAECGCR 429
```

RESULT 5
AAU14195
ID AAU14195 standard; Protein; 429 AA.
XX
AC AAU14195;
XX
DT 24-OCT-2001 (first entry)
XX
DE Human novel protein #66.
XX
KW Human; novel protein; Antianaemic; osteopathic; antiinflammatory;
KW immunomodulatory; cytostatic; neuroprotective; vulnerary; nootropic;
KW anticonvulsant; antiarthritic; cerebroprotective; antifungal; antiviral;
KW antibacterial; antiallergic; dermatological; haemostatic; antiasthmatic;
KW thrombolytic; immunogen; antibody; gene therapy; neurological disorder;
KW Parkinson's disease; inflammatory disorder; cancer; asthma; osteoporosis;
KW tissue regeneration; immune disorder.
XX
OS Homo sapiens.
XX
PN WO200155437-A2.
XX
PD 02-AUG-2001.
XX
PF 25-JAN-2001; 2001WO-US02623.
XX
PR 25-JAN-2000; 2000US-0491404.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Tang YT, Liu C, Drmanac RT;
XX
DR WPI; 2001-451939/48.
DR N-PSDB; AAS22500.
XX
PT Isolated polypeptides useful for treating anti-inflammatory diseases,
PT nervous system disorders, and for regenerating bone and cartilage -
XX
PS Example 4; Page 561-562; 894pp; English.
XX
CC The invention relates to polynucleotides encoding novel human
CC proteins or their active domains. The polypeptides, polynucleotides and
CC antibodies raised against the polypeptides are used in a method of
CC treatment of a mammal and prevention of disorders caused by the aberrant
CC protein expression or activity. The polypeptides can be used as
CC molecular weight markers, food supplements, and in antibody production.
CC The polypeptides are used to identify compounds which bind to the
CC polypeptides. Polynucleotides of the invention are used as probes and
CC primers, for sequencing, for chromosome or gene mapping, in the
CC production of recombinant proteins, and in generating anti-sense DNA or
CC RNA and in gene therapy. Polypeptides of the invention can be used to
CC target drugs to a tumour, in assays to determine biological activity, to
CC raise antibodies/elicit an immune response, to determine quantitative
CC protein levels, as tissue markers, and to isolate receptors or ligands.
CC Polypeptides of the invention may also be useful in treating platelet
CC disorders, stem cell disorders, regenerating bone, cartilage, tendon,
CC ligament and/or nerve tissue, wound healing, treating burns, promoting
CC the proliferation, differentiation and survival of stem cells, as a
CC contraceptive, treating osteoporosis and osteoarthritis, anaemia,
CC Alzheimer's, Parkinson's and Huntington's diseases, amylotrophic lateral
CC sclerosis, stroke, immune deficiencies resulting from bacterial, viral or
CC fungal infection or from autoimmunity, cancer, allergy, asthma,
CC graft-versus-host disease, eczema, haemophilia, thrombosis,
CC anti-inflammatory diseases, nervous system disorders, and infection.
CC The present sequence represents a protein of the invention.
XX
SQ Sequence 429 AA;

Query Match 83.8%; Score 1146.5; DB 22; Length 429;
Best Local Similarity 82.9%; Pred. No. 3.6e-109;
Matches 213; Conservative 19; Mismatches 24; Indels 1; Gaps 1;

```
Qy       1 DVLEDSETWDQATGTKTFLVSQDIRDEGWETLEVSSAVKRWVRADSTTNKNKLEVTVQSH 60
           |||: :: || || ||||||||||:||||||||||||||||||:||| :|||||||||:||
Db     174 DVLDGTDAWDSATETKTFLVSQDIQDEGWETLEVSSAVKRWVRSDSTKSKNKLEVTVESH 233

Qy      61 RESCDTLDISVPPGSKNLPFFVVFSNDRSNGTKETRLDLLKEMIGHEQETMLVKTAKNAY 120
           |: |||||||||||:||||||||||| |:|||||||: |:||| ||||::| | :|:
Db     234 RKGCDTLDISVPPGSRNLPFFVVFSNDHSSGTKETRLE-LREMISHEQESVLKKLSKDGS 292

Qy     121 QGAGESQEEEGLDGYTAVGPLLARRKRSTGASSHCQKTSLRVNFEDIGWDSWIIAPKEYD 180
             |||| || ||: | |  ||||||| || ||||||||||||||||||||||||||||:
Db     293 TEAGESSHEEDTDGHVAAGSTLARRKRSAGAGSHCQKTSLRVNFEDIGWDSWIIAPKEYE 352

Qy     181 AYECKGGCFPPLADDVTPTKHAIVQTLVHLKFPTKVGKACCVPTKLSPISILYKDDMGVP 240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||:|||||||||
Db     353 AYECKGGCFPPLADDVTPTKHAIVQTLVHLKFPTKVGKACCVPTKLSPISVLYKDDMGVP 412

Qy     241 TLKYHYEGMSVAECGCR 257
           |||||||||||||||||
Db     413 TLKYHYEGMSVAECGCR 429
```

RESULT 6
AAU14431
ID AAU14431 standard; Protein; 429 AA.
XX
AC AAU14431;
XX
DT 24-OCT-2001 (first entry)
XX
DE Human novel protein #302.
XX
KW Human; novel protein; Antianaemic; osteopathic; antiinflammatory;
KW immunomodulatory; cytostatic; neuroprotective; vulnerary; nootropic;
KW anticonvulsant; antiarthritic; cerebroprotective; antifungal; antiviral;
KW antibacterial; antiallergic; dermatological; haemostatic; antiasthmatic;
KW thrombolytic; immunogen; antibody; gene therapy; neurological disorder;
KW Parkinson's disease; inflammatory disorder; cancer; asthma; osteoporosis;
KW tissue regeneration; immune disorder.
XX
OS Homo sapiens.
XX
PN WO200155437-A2.
XX
PD 02-AUG-2001.
XX
PF 25-JAN-2001; 2001WO-US02623.
XX
PR 25-JAN-2000; 2000US-0491404.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Tang YT, Liu C, Drmanac RT;

```
XX
DR   WPI; 2001-451939/48.
DR   N-PSDB; AAS22736.
XX
PT   Isolated polypeptides useful for treating anti-inflammatory diseases,
PT   nervous system disorders, and for regenerating bone and cartilage -
XX
PS   Example 4; Page 812-813; 894pp; English.
XX
CC   The invention relates to polynucleotides encoding novel human
CC   proteins or their active domains. The polypeptides, polynucleotides and
CC   antibodies raised against the polypeptides are used in a method of
CC   treatment of a mammal and prevention of disorders caused by the aberrant
CC   protein expression or activity.  The polypeptides can be used as
CC   molecular weight markers, food supplements, and in antibody production.
CC   The polypeptides are used to identify compounds which bind to the
CC   polypeptides. Polynucleotides of the invention are used as probes and
CC   primers, for sequencing, for chromosome or gene mapping, in the
CC   production of recombinant proteins, and in generating anti-sense DNA or
CC   RNA and in gene therapy.  Polypeptides of the invention can be used to
CC   target drugs to a tumour, in assays to determine biological activity, to
CC   raise antibodies/elicit an immune response, to determine quantitative
CC   protein levels, as tissue markers, and to isolate receptors or ligands.
CC   Polypeptides of the invention may also be useful in treating platelet
CC   disorders, stem cell disorders, regenerating bone, cartilage, tendon,
CC   ligament and/or nerve tissue, wound healing, treating burns, promoting
CC   the proliferation, differentiation and survival of stem cells, as a
CC   contraceptive, treating osteoporosis and osteoarthritis, anaemia,
CC   Alzheimer's, Parkinson's and Huntington's diseases, amylotrophic lateral
CC   sclerosis, stroke, immune deficiencies resulting from bacterial, viral or
CC   fungal infection or from autoimmunity, cancer, allergy, asthma,
CC   graft-versus-host disease, eczema, haemophilia, thrombosis,
CC   anti-inflammatory diseases, nervous system disorders, and infection.
CC   The present sequence represents a protein of the invention.
XX
SQ   Sequence   429 AA;

  Query Match             83.8%;  Score 1146.5;  DB 22;  Length 429;
  Best Local Similarity   82.9%;  Pred. No. 3.6e-109;
  Matches  213;  Conservative   19;  Mismatches   24;  Indels    1;  Gaps    1;

Qy       1 DVLEDSETWDQATGTKTFLVSQDIRDEGWETLEVSSAVKRWVRADSTTNKNKLEVTVQSH 60
           |||: :: || || ||||||||||:|||||||||||||||||||:||| :|||||||:||
Db     174 DVLDGTDAWDSATETKTFLVSQDIQDEGWETLEVSSAVKRWVRSDSTKSKNKLEVTVESH 233

Qy      61 RESCDTLDISVPPGSKNLPFFVVFSNDRSNGTKETRLDLLKEMIGHEQETMLVKTAKNAY 120
           |: ||||||||||||:||||||||||| |:|||||||: |:||| ||||::| | :|:
Db     234 RKGCDTLDISVPPGSRNLPFFVVFSNDHSSGTKETRLE-LREMISHEQESVLKKLSKDGS 292

Qy     121 QGAGESQEEEGLDGYTAVGPLLARRKRSTGASSHCQKTSLRVNFEDIGWDSWIIAPKEYD 180
             ||||  ||  ||: | |  |||||||| || ||||||||||||||||||||||||||:
Db     293 TEAGESSHEEDTDGHVAAGSTLARRKRSAGAGSHCQKTSLRVNFEDIGWDSWIIAPKEYE 352

Qy     181 AYECKGGCFFPLADDVTPTKHAIVQTLVHLKFPTKVGKACCVPTKLSPISILYKDDMGVP 240
           ||||||||||||||||||||||||||||||||||||||||||||||:|||||||||
Db     353 AYECKGGCFFPLADDVTPTKHAIVQTLVHLKFPTKVGKACCVPTKLSPISVLYKDDMGVP 412

Qy     241 TLKYHYEGMSVAECGCR 257
           |||||||||||||||||
Db     413 TLKYHYEGMSVAECGCR 429


RESULT 7
AAR68607
ID   AAR68607 standard; Protein; 427 AA.
XX
AC   AAR68607;
XX
DT   02-AUG-1995  (first entry)
XX
DE   Dorsalin I.
XX
KW   Dorsalin; neural crest cell differentiation; nerve cells; neurons;
KW   regeneration; wound healing; bone growth; tumour; neurofibroma;
KW   Schwann cell tumour.
XX
OS   Gallus gallus.
XX
PN   WO9428016-A.
XX
PD   08-DEC-1994.
XX
PF   20-MAY-1994;   94WO-US05743.
XX
PR   20-MAY-1993;   93US-0065844.
XX
PA   (UYCO ) UNIV COLUMBIA NEW YORK.
XX
PI   Basler K,  Jessel TM,  Yamada T;
XX
DR   WPI; 1995-022710/03.
DR   N-PSDB; AAQ80276.
XX
PT   New nucleic acid encoding vertebrate dorsalin-1 - and related
PT   vectors, transformed cells, proteins and antibodies, used e.g. to
PT   regenerate nerve cells, promote bone growth etc.
XX
PS   Claim 17; Page 66-67; 99pp; English.
XX
CC   Dorsalin-I is able to stimulate neural crest cell differentiation in
CC   culture.  It can be used in vivo to regenerate nerve cells; to
CC   promote bone growth and wound healing; and to treat (inhibit growth
CC   of) neural tumours (specifically neurofibroma or Schwann cell
CC   tumours).  Fragments of the coding sequence can be useful as probes
CC   for studying nerve development; for isolating other dorsalin
CC   encoding genes and to locate tissues expressing these genes.
XX
SQ   Sequence   427 AA;

  Query Match             59.1%;  Score 808.5;  DB 16;  Length 427;
  Best Local Similarity   61.9%;  Pred. No. 2.2e-74;
  Matches  161;  Conservative   31;  Mismatches   59;  Indels    9;  Gaps    5;

Qy       1 DVLEDSETWDQATGTKTFLVSQDIRDEGWETLEVSSAVKRWVRADSTTNKNKLEVTVQSH 60
           ||| | : |:    ||: |||  |:| |||  ||||||||||:||    |||||| ::|
Db     174 DVL-DGDHWENKESTKSLLVSHSIQDCGWEMFEVSSAVKRWVKADKMKTKNKLEVVIESK 232

Qy      61 RES---CDTLDISVPPGSKNLPFFVVFSNDRSNGTKETRLDLLKEMIGHEQETMLVKTAK 117
             |   |  |||:|   :|||| :|||||||||||||::: |:||| ||||::| |  |
Db     233 DLSGFPCGKLDITVTHDTKNLPLLIVFSNDRSNGTKETKVE-LREMIVHEQESVLNKLGK 291

Qy     118 NAYQGAGESQEEEGLDGYTAVGPLLARRKRSTGASSHCQKTSLRVNFEDIGWDSWIIAPK 177
           |      | :||: :          :| ||| || :||::||| |||::|||||||||
Db     292 NDSSSEEEQREEKAI---ARPRQHSSRSKRSIGA-NHCRRTSLHVNFKEIGWDSWIIAPK 347

Qy     178 EYDAYECKGGCFFPLADDVTPTKHAIVQTLVHLKFPTKVGKACCVPTKLSPISILYKDDM 237
           :|:|:|||||||||| |:|||||||||||||: | |  |||||||| |||||||
Db     348 DYEAFECKGGCFFPLTDNVTPTKHAIVQTLVHLQNPKKASKACCVPTKLDAISILYKDDA 407

Qy     238 GVPTLKYHYEGMSVAECGCR 257
           ||||| |:|||| |||||||
Db     408 GVPTLIYNYEGMKVAECGCR 427


RESULT 8
AAR86903
ID   AAR86903 standard; Protein; 150 AA.
XX
AC   AAR86903;
XX
DT   10-MAY-1996  (first entry)
XX
DE   Human BMP-9 polypeptide.
XX
KW   BMP-9; bone; cartilage; wound healing; liver; tissue repair.
XX
```

```
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   Cleavage-site   40..41
FT                   /note= "proteolytic cleavage site"
FT   Protein         41..150
FT                   /label= Mat_protein
XX
PN   WO9533830-A1.
XX
PD   14-DEC-1995.
XX
PF   05-JUN-1995;   95WO-US07084.
XX
PR   06-JUN-1994;   94US-0254353.
XX
PA   (GEMY ) GENETICS INST INC.
XX
PI   Celeste AJ,  Rosen VA,  Song JJ,  Thies SR,  Wozney JM;
XX
DR   WPI; 1996-040235/04.
DR   N-PSDB; AAT13143.
XX
PT   BMP-9 protein able to induce bone and cartilage formation - also
PT   useful in wound healing, tissue repair and for inducing hepatic
PT   growth and function
XX
PS   Claim 1; Page 45-46; 75pp; English.
XX
CC   Human BMP-9 polypeptide (AAR86903) is the product of a DNA fragment
CC   (AAQ74084) isolated from a human genomic DNA library.  Cleavage of
CC   the polypeptide at the consensus proteolytic processing site
CC   gives a 110-amino acid mature protein that dimerizes to form a
CC   homodimeric, active species.   Recombinant BMP-9 can be obtd. by
CC   expression of the genomic DNA in procaryotic or eucaryotic hosts.
CC   It has therapeutic applns. in the formation of bone and
CC   cartilage, wound healing and tissue repair.
XX
SQ   Sequence   150 AA;

  Query Match             46.3%;  Score 634;  DB 17;  Length 150;
  Best Local Similarity   92.7%;  Pred. No. 4.4e-57;
  Matches  114;  Conservative    3;  Mismatches    6;  Indels    0;  Gaps    0;

Qy     135 YTAVGPLLARRKRSTGASSHCQKTSLRVNFEDIGWDSWIIAPKEYDAYECKGGCFFPLAD 194
           : | |  ||||||| || |||||||||||||||||||||||||:|||||||||||||||
Db      28 HVAAGSTLARRKRSAGAGSHCQKTSLRVNFEDIGWDSWIIAPKEYEAYECKGGCFFPLAD 87

Qy     195 DVTPTKHAIVQTLVHLKFPTKVGKACCVPTKLSPISILYKDDMGVPTLKYHYEGMSVAEC 254
           ||||||||||||||||||||||||||||||||||||:||||||||||||||||||||||
Db      88 DVTPTKHAIVQTLVHLKFPTKVGKACCVPTKLSPISVLYKDDMGVPTLKYHYEGMSVAEC 147

Qy     255 GCR 257
           |||
Db     148 GCR 150


RESULT 9
AAY82272
ID   AAY82272 standard; Protein; 150 AA.
XX
AC   AAY82272;
XX
DT   16-JUN-2000  (first entry)
XX
DE   Human bone morphogenetic protein 9 SEQ ID NO:9.
XX
KW   Bone morphogenetic protein 9; BMP-9; bone defect; cartilage defect;
KW   wound healing; tissue repair; hepatic growth; antiinfertility;
KW   osteopathic; antiarthritic; vulnerary.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   Peptide         1..40
FT                   /label= signal
FT   Protein         41..150
FT                   /label= Bone_morphogenetic_protein-9
XX
PN   US6034062-A.
XX
PD   07-MAR-2000.
XX
PF   13-MAR-1997;   97US-0815652.
XX
PR   13-MAR-1997;   97US-0815652.
XX
PA   (GEMY ) GENETICS INST INC.
XX
PI   Thies RS,  Song JJ;
XX
DR   WPI; 2000-255707/22.
DR   N-PSDB; AAZ95747.
XX
PT   Novel bone-morphogenetic protein (BMP)-9 polypeptides, useful in the
PT   treatment of bone and cartilage defects and in wound healing and tissue
PT   repair, as well as hepatic growth and function  -
XX
PS   Claim 1; Column 45-46; 36pp; English.
XX
CC   A method has been developed for decreasing proliferation of kidney or
CC   lung epithelial cells, or for inhibiting kidney tubule formation, in a
CC   patient. The method comprises administering an amount of a bone
CC   morphogenetic protein 9 (BMP-9) protein. BMP-9 polypeptides from the
CC   invention may be used to induce bone and cartilage formation or other
CC   connective tissue formation, in periodontal disease, in wound healing
CC   and tissue repair, hepatic growth and function, and in regulating lung
CC   and kidney cell function. The polypeptides can be used for healing of
CC   bone fractures, open fracture reduction, improved fixation of artificial
CC   joints, tooth repair processes, treatment of osteoporosis and
CC   osteoarthritis, and as a dietary supplement or as a component of cell
CC   culture media. BMP-9 may also be used for repair and regenerating of
CC   liver cells. The BMP-9 polypeptides may also have angiogenic,
CC   chemotactic and/or chemoattractant properties, and may induce collagen
CC   synthesis, fibrosis, differentiation processes, cell proliferative
CC   responses, cell adhesion responses, and migration. When dimerized, the
CC   BMP-9 polypeptide is expected to demonstrate effects on the production
CC   of follicle stimulating hormone (FSH), and so may be used as a
CC   contraceptive, as a fertility inducing therapeutic, and for advancing
CC   the onset of fertility in sexually immature mammals, so as to increase
CC   the lifetime reproductive performance of domestic animals. The present
CC   sequence represents human BMP-9.
XX
SQ   Sequence   150 AA;

  Query Match             46.3%;  Score 634;  DB 21;  Length 150;
  Best Local Similarity   92.7%;  Pred. No. 4.4e-57;
  Matches  114;  Conservative    3;  Mismatches    6;  Indels    0;  Gaps    0;

Qy     135 YTAVGPLLARRKRSTGASSHCQKTSLRVNFEDIGWDSWIIAPKEYDAYECKGGCFFPLAD 194
           : | |  ||||||| || |||||||||||||||||||||||||:|||||||||||||||
Db      28 HVAAGSTLARRKRSAGAGSHCQKTSLRVNFEDIGWDSWIIAPKEYEAYECKGGCFFPLAD 87

Qy     195 DVTPTKHAIVQTLVHLKFPTKVGKACCVPTKLSPISILYKDDMGVPTLKYHYEGMSVAEC 254
           ||||||||||||||||||||||||||||||||||||:||||||||||||||||||||||
Db      88 DVTPTKHAIVQTLVHLKFPTKVGKACCVPTKLSPISVLYKDDMGVPTLKYHYEGMSVAEC 147

Qy     255 GCR 257
           |||
Db     148 GCR 150


RESULT 10
AAR31374
ID   AAR31374 standard; Protein; 150 AA.
XX
```

```
AC   AAR31374;
XX
DT   24-JUN-1993  (first entry)
XX
DE   Human bone morphogenic protein-9 (BMP-9).
XX
KW   bone formation; cartilage formation; wound healing; tissue repair;
KW   surgery; fracture treatment; periodontal disease; osteoporosis;
KW   increase neuronal survival; transplantation; nerves; nervous system.
XX
OS   Homo sapiens.
XX
PN   WO9300432-A.
XX
PD   07-JAN-1993.
XX
PF   25-JUN-1992;   92WO-US05374.
XX
PR   25-JUN-1991;   91US-0720590.
XX
PA   (GEMY ) GENETICS INST INC.
XX
PI   Celeste AJ,  Wozney JM;
XX
DR   WPI; 1993-036379/04.
DR   N-PSDB; AAQ35244.
XX
PT   New BMP-9 polypeptide(s) induce bone and cartilage formation -
PT   used for treating osteoporosis and fractures, healing wounds and
PT   increasing neuronal survival
XX
PS   Example 2; Fig 3; 60pp; English.
XX
CC   This sequence is human bone morphogenic protein-9. The protein can
CC   be used to induce bone and/or cartilage formation and in wound healing
CC   and tissue repair. It can be used in surgery of the treatment of
CC   fractures, periodontal disease or osteoporosis. The protein can also
CC   increase neuronal survival and can be used in transplantation and
CC   treatment of conditions exhibiting a decrease in neuronal survival.
XX
SQ   Sequence   150 AA;

  Query Match             45.2%;  Score 619;  DB 14;  Length 150;
  Best Local Similarity   91.1%;  Pred. No. 1.5e-55;
  Matches  112;  Conservative    2;  Mismatches    9;  Indels    0;  Gaps    0;

Qy     135 YTAVGPLLARRKRSTGASSHCQKTSLRVNFEDIGWDSWIIAPKEYDAYECKGGCFFPLAD 194
           : | |  ||||||| || ||||||||||||||||||||||||||||| ||||||||||||
Db      28 HVAAGSTLARRKRSAGAGSHCQKTSLRVNFEDIGWDSWIIAPKEYKAYECKGGCFFPLAD 87

Qy     195 DVTPTKHAIVQTLVHLKFPTKVGKACCVPTKLSPISILYKDDMGVPTLKYHYEGMSVAEC 254
           |||||||||||||||||||||||||||| ||| ||||:||||||||||||||||||||||||
Db      88 DVTPTKHAIVQTLVHLKFPTKVGKACCEPTKQSPISVLYKDDMGVPTLKYHYEGMSVAEC 147

Qy     255 GCR 257
           |||
Db     148 GCR 150


RESULT 11
AAM51930
ID   AAM51930 standard; protein; 110 AA.
XX
AC   AAM51930;
XX
DT   01-FEB-2002  (first entry)
XX
DE   Human TGFbeta protein superfamily protein GDF2.
XX
KW   Human; TGFbeta; transforming growth factor beta; mutant; antagonist;
KW   agonist; ectopic bone formation; psoriasis; muscular atrophy; scar;
KW   formation; fibrosis; cirrhosis; osteopathic; antipsoriatic;
KW   antifibrotic; hepatotropic; vulnerary; GDF2.
XX
OS   Homo sapiens.
XX
PN   DE10026713-A1.
XX
PD   06-DEC-2001.
XX
PF   30-MAY-2000; 2000DE-1026713.
XX
PR   30-MAY-2000; 2000DE-1026713.
XX
PA   (SEBA/) SEBALD W.
XX
PI   Sebald W,  Nickel J;
XX
DR   WPI; 2002-042559/06.
XX
PT   New mutein of transforming growth factor-beta superfamily protein,
PT   useful for treating or preventing e.g. ectopic bone formation, competes
PT   for receptor binding  -
XX
PS   Disclosure; Fig 6; 54pp; German.
XX
CC   The present invention relates to muteins of a chain of a protein which,
CC   when in the form of a homodimer, has antagonistic or partial agonistic
CC   activity, and where the mutation results in the protein binding with low
CC   affinity to its receptor. The protein is a member of the transforming
CC   growth factor beta (TGFbeta) superfamily. The mutant sequences of the
CC   invention can be used in the treatment of diseases associated with the
CC   overexpression of TGFbeta family proteins, including ectopic bone
CC   formation, psoriasis, muscular atrophy, scar formation, fibrosis and
CC   cirrhosis. The present sequence is the human GDF2 protein.
XX
SQ   Sequence   110 AA;

  Query Match             43.6%;  Score 596;  DB 23;  Length 110;
  Best Local Similarity   96.4%;  Pred. No. 2.3e-53;
  Matches  106;  Conservative    2;  Mismatches    2;  Indels    0;  Gaps    0;

Qy     148 STGASSHCQKTSLRVNFEDIGWDSWIIAPKEYDAYECKGGCFFPLADDVTPTKHAIVQTL 207
           | || |||||||||||||||||||||||||:||||||||||||||||||||||||||||
Db       1 SAGAGSHCQKTSLRVNFEDIGWDSWIIAPKEYEAYECKGGCFFPLADDVTPTKHAIVQTL 60

Qy     208 VHLKFPTKVGKACCVPTKLSPISILYKDDMGVPTLKYHYEGMSVAECGCR 257
           |||||||||||||||||||||||:|||||||||||||||||||||||||||
Db      61 VHLKFPTKVGKACCVPTKLSPISVLYKDDMGVPTLKYHYEGMSVAECGCR 110


RESULT 12
AAR66200
ID   AAR66200 standard; Protein; 312 AA.
XX
AC   AAR66200;
XX
DT   06-AUG-1995  (first entry)
XX
DE   Part of bovine bone morphogenetic protein (BMP)-10.
XX
KW   Bone morphogenetic protein; BMP; bone formation; cartilage;
KW   peridontal disease; osteoporosis.
XX
OS   Bos taurus.
XX
FH   Key             Location/Qualifiers
FT   Peptide         1..204
FT                   /label= part of propeptide
FT   Protein         205..312
FT                   /label= mature peptide
XX
PN   WO9426893-A.
XX
PD   24-NOV-1994.
XX
```

```
PF   12-MAY-1994;   94WO-US05290.
XX
PR   12-MAY-1993;   93US-0061695.
XX
PA   (GEMY ) GENETICS INST INC.
XX
PI   Celeste AJ,  Wozney JM;
XX
DR   WPI; 1995-006789/01.
DR   N-PSDB; AAQ79524.
XX
PT   New bone morphogenetic protein 10 and related nucleic acid -
PT   vectors and transformed cells, induces formation of cartilage and
PT   bone, useful for healing fractures, wounds, etc., or treating
PT   osteoporosis.
XX
PS   Disclosure; Page 34-35; 51pp; English.
XX
CC   BMP-10 is produced by culturing a host cell transformed with a DNA
CC   sequence comprising AAQ79524 from bps 167 to 1102 or from  bps 779-
CC   1102. A DNA sequence comprising bps 779 or 797-1102 of AAQ79524 is
CC   claimed. AAQ79524 is from clone lambda 7r-20 which was derived as
CC   follows. A phage library of bovine genomic DNA was screened (a)
CC   under low stringency conditions with a human BMP-7 cDNA fragment and
CC   (b) under high stringency conditions with BMP-5, -6 and -7 probes.
CC   Once clone positive in the first screen but negative in the second,
CC   lambda 7r-20 (ATCC 75452), was sequenced. It encodes at least of
CC   a portion of bovine BMP-10. The processing of BMP-10 into the mature
CC   form is expected to involve dimerisation and removal of the
CC   N-terminal region. BMP-10 may exist as a homodimer with both
CC   subunits being the 108 AA mature peptides. Alternatively, it may be
CC   a heterodimer with one 108 AA subunit and the other subunit from
CC   any one of BMP 1-9. Fragments derived from clone lambda 7r-20 may
CC   be used to isolate homologous human DNA (see AAQ79525/R66201, AAQ79526,
CC   AAQ79527, AAQ79531/R66202).
XX
SQ   Sequence   312 AA;

  Query Match             38.0%;  Score 519.5;  DB 16;  Length 312;
  Best Local Similarity   42.8%;  Pred. No. 7.6e-45;
  Matches  119;  Conservative   43;  Mismatches   73;  Indels   43;  Gaps   10;

Qy       1 DVLEDSETWDQATGTKTFLVSQDI--RDEGWETLEVSSAVKRWVRADSTTNKNKLEVTVQ 58
Db      57 EVLESKED-HEGERNMLVLVSGEIYGTNSEWETFDVTDAIRHWQKSGSST--HQLEVHIE 113

Qy      59 SHRESCDT-----LDISVPPGSKNLPFFVVFSNDRSNGTKETRLDLLKEMIGHEQETMLV 113
Db     114 SKHEMEDTLGRGQLEIDTSARNKHDPLLVVFSDDQS--SEKERKEELDEMIAHEQ----- 166

Qy     114 KTAKNAYQGAGESQEEEGLDGYTAVGP----LLARRK----------RSTGASSHCQKTS 159
Db     167 ----------FPEMDNLDLDGYSN-GPGEEALLQMRSNIIYDSTARIRRNAKGNYCKRTP 215

Qy     160 LRVNFEDIGWDSWIIAPKEYDAYECKGGCFFPLADDVTPTKHAIVQTLVHLKFPTKVGKA 219
Db     216 LYIDFKEIGWDSWIIAPPGYEAYECRGVCNYPLAEHLTPTKHAIIQALVHLKNSQKASKA 275

Qy     220 CCVPTKLSPISILYKDDMGVPTLKYHYEGMSVAECGCR 257
Db     276 CCVPTKLEPISILYL-DKGVVTYKFKYEGMAVSECGCR 312


RESULT 13
AAR66202
ID   AAR66202 standard; Protein; 424 AA.
XX
AC   AAR66202;
XX
DT   06-AUG-1995  (first entry)
XX
DE   Complete human bone morphogenetic protein (BMP)-10.
XX
KW   Bone morphogenetic protein; BMP; bone formation; cartilage;
KW   peridontal disease; osteoporosis; primer.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   Protein         317..424
FT                   /label= mature protein
XX
PN   WO9426893-A.
XX
PD   24-NOV-1994.
XX
PF   12-MAY-1994;   94WO-US05290.
XX
PR   12-MAY-1993;   93US-0061695.
XX
PA   (GEMY ) GENETICS INST INC.
XX
PI   Celeste AJ,  Wozney JM;
XX
DR   WPI; 1995-006789/01.
DR   N-PSDB; AAQ79531.
XX
PT   New bone morphogenetic protein 10 and related nucleic acid -
PT   vectors and transformed cells, induces formation of cartilage and
PT   bone, useful for healing fractures, wounds, etc., or treating
PT   osteoporosis.
XX
PS   Claim 18; Page 39-41; 51pp; English.
XX
CC   AAQ79525/R66201 comprise the partial human BMP-10 sequence.
CC   An oligo probe based on nts 85-114 of AAQ79525 was
CC   used to screen a human fetal liver cDNA library. The cDNA of one
CC   of the positively hybridising recombinants, named HFL-3, was
CC   isolated and deposited at the ATCC. A portion of the DNA sequence
CC   in clone HFL-3 is set out in AAQ79531/R66202. An oligo probe based
CC   on nts 355-384 of AAQ79531 was used to screen a human genomic
CC   library. The DNA from one of the positively hybridising recombinants,
CC   named 20GEN-3, was isolated and deposited at the ATCC, and a
CC   portion of this clone is set forth in AAQ79531. A portion of
CC   20GEN.3 was determined to be identical to cDNA clone HFL-3.
CC   The extent of this overlap is nts 219-316. Nts 1-218 are derived
CC   from genomic clone 20GEN.3 and nts 317-1584 are derived from cDNA
CC   clone HFL-3. Nts 1108 or 1126-1431 are claimed. The purified
CC   mature human BMP-10 polypeptide is also claimed.
XX
SQ   Sequence   424 AA;

  Query Match             37.8%;  Score 516.5;  DB 16;  Length 424;
  Best Local Similarity   42.8%;  Pred. No. 2.4e-44;
  Matches  119;  Conservative   48;  Mismatches   68;  Indels   43;  Gaps   11;

Qy       1 DVLEDSETWDQATGTKTFLVSQDI--RDEGWETLEVSSAVKRWVRADSTTNKNKLEVTVQ 58
Db     169 EVLE-SKGDNEGERNMLVLVSGEIYGTNSEWETFDVTDAIRRWQKSGSST--HQLEVHIE 225

Qy      59 S-HRESCDT----LDISVPPGSKNLPFFVVFSNDRSNGTKETRLDLLKEMIGHEQETMLV 113
Db     226 SKHDEAEDASSGRLEIDTSAQNKHNPLLIVFSDDQS--SDKERKEELNEMISHEQLPEL- 282

Qy     114 KTAKNAYQGAGESQEEEGLDGYTAVGP----LLARRK----------RSTGASSHCQKTS 159
Db     283 -------------DNLGLDSFSS-GPGEEALLQMRSNIIYDSTARIRRNAKGNYCKRTP 327

Qy     160 LRVNFEDIGWDSWIIAPKEYDAYECKGGCFFPLADDVTPTKHAIVQTLVHLKFPTKVGKA 219
Db     328 LYIDFKEIGWDSWIIAPPGYEAYECRGVCNYPLAEHLTPTKHAIIQALVHLKNSQKASKA 387

Qy     220 CCVPTKLSPISILYKDDMGVPTLKYHYEGMSVAECGCR 257
Db     388 CCVPTKLEPISILYL-DKGVVTYKFKYEGMAVSECGCR 424
```

```
RESULT 14
AAY92029
ID   AAY92029 standard; Protein; 424 AA.
XX
AC   AAY92029;
XX
DT   19-JUL-2000  (first entry)
XX
DE   Human bone morphogenic protein-10 (BMP-10).
XX
KW   human bone morphogenic protein-10; BMP-10; Vgrl growth factor; CKGF;
KW   mutant; cystine knot growth factor; hairpin loop; infertility.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   Misc-difference 1..326
FT                   /note= "optionally mutated to increase electrostatic
FT                          interaction between beta hairpin structure and
FT                          a receptor"
FT   Domain          327..353
FT                   /label= beta_hairpin_loop_1
FT                   /note= "mutant optionally comprises one or more
FT                          substitutions in these residues"
FT   Misc-difference 354..392
FT                   /note= "optionally mutated to increase electrostatic
FT                          interaction between beta hairpin structure and
FT                          a receptor"
FT   Domain          394..416
FT                   /label= beta_hairpin_loop_3
FT                   /note= "mutant optionally comprises one or more
FT                          substitutions in these residues"
FT   Misc-difference 417..424
FT                   /note= "optionally mutated to increase electrostatic
FT                          interaction between beta hairpin structure and
FT                          a receptor"
XX
PN   WO200017360-A1.
XX
PD   30-MAR-2000.
XX
PF   19-MAR-1999;   99WO-US05908.
XX
PR   22-SEP-1998;   98WO-US19772.
XX
PA   (UYMA-) UNIV MARYLAND BALTIMORE.
XX
PI   Weintraub BD,  Szkudlinski MW;
XX
DR   WPI; 2000-283585/24.
XX
PT   New mutant cystine knot growth factor proteins comprising one or more
PT   mutant subunits, useful for treating or preventing diseases e.g.
PT   hypothyroidism and thyroid cancer
XX
PS   Claim 475; Page 310; 320pp; English.
XX
CC   This is the wild type human bone morphogenic protein-10 (BMP-10).
CC   Mutants comprise at least one electrostatic charge altering mutation in a
CC   beta hairpin loop, resulting in increased bioactivity.
CC   Mutant cystine knot growth factor (CKGF) proteins comprising one or more
CC   mutant subunits and having novel properties or improved pharmacological
CC   properties, compared to wild type CKGFs, are claimed.  The CKGF
CC   superfamily comprises at least four families of growth factors: the
CC   glycoprotein hormones, the platelet-derived growth factor (PDGF) family,
CC   the neurotrophins and the transforming growth factor-beta family; the
CC   families are known to be structurally similar (especially comprising the
CC   cystine knot topology) and it was shown that mutations at certain
CC   positions in the CKGF hairpin loops of family members and other members
CC   of the CKGF superfamily could significantly alter the biological
CC   activities of the CKGF.
CC   Mutant transforming growth factor family proteins or analogues are useful
CC   for treatment of ovulatory dysfunction, luteal phase defect, unexplained
CC   infertility, time-limited conception and in assisted reproduction.
XX
SQ   Sequence   424 AA;

  Query Match             37.8%;  Score 516.5;  DB 21;  Length 424;
  Best Local Similarity   42.8%;  Pred. No. 2.4e-44;
  Matches  119;  Conservative   48;  Mismatches   68;  Indels   43;  Gaps   11;

Qy       1 DVLEDSETWDQATGTKTFLVSQDI--RDEGWETLEVSSAVKRWVRADSTTNKNKLEVTVQ 58
           :||| |:  ::       ||| :|   :  ||| :|: |::|| :: |:|  ::||| ::
Db     169 EVLE-SKGDNEGERNMLVLVSGEIYGTNSEWETFDVTDAIRRWQKSGSST--HQLEVHIE 225

Qy      59 S-HRESCDT----LDISVPPGSKNLPFFVVFSNDRSNGTKETRLDLLKEMIGHEQETMLV 113
           | | |: |     |:|     :|: |  :|||:|:|  : : | : | ||| |||   |
Db     226 SKHDEAEDASSGRLEIDTSAQNKHNPLLIVFSDDQS--SDKERKEELNEMISHEQLPEL- 282

Qy     114 KTAKNAYQGAGESQEEEGLDGYTAVGP----LLARRK----------RSTGASSHCQKTS 159
                       :  ||| ::: ||    ||  |          |     ::|::|
Db     283 --------------DNLGLDSFSS-GPGEEALLQMRSNIIYDSTARIRRNAKGNYCKRTP 327

Qy     160 LRVNFEDIGWDSWIIAPKEYDAYECKGGCFFPLADDVTPTKHAIVQTLVHLKFPTKVGKA 219
           | ::|::||||||||||  |:|||:| | :|||: :|||||||:| |||||   |  ||
Db     328 LYIDFKEIGWDSWIIAPPGYEAYECRGVCNYPLAEHLTPTKHAIIQALVHLKNSQKASKA 387

Qy     220 CCVPTKLSPISILYKDDMGVPTLKYHYEGMSVAECGCR 257
           ||||||| ||||||  | || | |: ||||:|:|||||
Db     388 CCVPTKLEPISILYL-DKGVVTYKFKYEGMAVSECGCR 424


RESULT 15
AAY06304
ID   AAY06304 standard; Protein; 421 AA.
XX
AC   AAY06304;
XX
DT   31-AUG-1999  (first entry)
XX
DE   Mouse pancreatic-derived factor PDF1.
XX
KW   PDF1; pancreatic-derived factor; mouse; signal transduction;
KW   transforming growth factor-beta; pancreas; diabetes; therapy;
KW   diagnosis.
XX
OS   Mus musculus.
XX
FH   Key             Location/Qualifiers
FT   Domain          320..333
FT                   /note= "core motif"
XX
PN   WO9929719-A2.
XX
PD   17-JUN-1999.
XX
PF   09-DEC-1998;   98WO-US26165.
XX
PR   09-DEC-1997;   97US-0069071.
XX
PA   (ONTO-) ONTOGENY INC.
XX
PI   Edlund H;
XX
DR   WPI; 1999-385571/32.
DR   N-PSDB; AAX59115.
XX
PT   New isolated pancreatic-derived factor polypeptides used for the
PT   treatment of a pancreatic disorder
XX
PS   Claim 2; Page 82-84; 84pp; English.
XX
CC   This sequence represents novel mouse pancreatic-derived factor
CC   PDF1, as predicted from a cDNA clone (see AAX59115) obtained from
CC   foetal pancreatic bud cDNA.  PDF1 shows homology to members of the
```

CC transforming growth factor-beta superfamily, especially BMP2a.
CC Its expression in the developing pancreas indicates a role in the
CC development of endodermal tissue and in particular in pancreas
CC development. Experimental evidence indicates a functional role
CC in signal transduction mediated by members of the TGF-beta
CC superfamily. PDF1 and PDF2 (see AAY06303) play a role in determining
CC tissue fate and in maintenance of differentiated states. They can
CC be used for the manufacture of medicaments for the treatment of a
CC pancreatic disorder or to modulate growth and/or differentiation of
CC pancreatic cells or stem cells capable of differentiating to
CC pancreatic cells (claimed). PDF polynucleotides, polypeptides and
CC modulator compounds can be used for treating e.g. diabetes, liver
CC disease, various cancers, leukemia, psoriasis, bone disease,
CC fibroproliferative disorders, atherosclerosis, chronic inflammation,
CC Alzheimer's disease, Parkinson's disease, Huntington's chorea,
CC amylotrophic lateral sclerosis, spinocerebellar degeneration,
CC disorders associated with connective tissue, e.g. de-differentiation
CC of chondrocytes or osteocytes, vascular disorders which involve
CC de-differentiation of endothelial tissue and smooth muscle cells,
CC gastric ulcers characterized by degenerative changes in glandular
CC cells, and renal conditions marked by failure to differentiate,
CC e.g. Wilm's tumors, to reform injured tissue, to improve grafting
CC and morphology of transplanted tissue, e.g. PDF agonists and
CC antagonists can be used in a differential manner to regulate
CC different stages of organ repair after physical, chemical or
CC pathological insult, in repair of pancreatic endocrine or exocrine
CC function, cartilage repair, increasing bone density, liver repair
CC subsequent to a partial hepatectomy, or to promote regeneration of
CC lung tissue in the treatment of emphysema. Modulation of the
CC function of PDF can be used in both cell culture and therapeutic
CC methods involving generation and maintenance of beta-cells and
CC possibly also for non-pancreatic tissue, such as in controlling the
CC development and maintenance of tissue from the digestive tract,
CC spleen, lungs, and other organs which derive from the primitive
CC gut. The products can also be used for treating endocrinopathies,
CC sepsis (including malaria), congestive cardiac failure, hepatic and
CC renal insufficiences, various genetic abnormalities of metabolism,
CC and exogenous toxins (such as alcohol). PDF therapeutics can also
CC be used in the treatment of hyperproliferative vascular disorders,
CC e.g. smooth muscle hyperplasia (such as atherosclerosis) or
CC restenosis, as well as other disorders characterized by fibrosis,
CC e.g. rheumatoid arthritis, insulin dependent diabetes mellitus,
CC glomerulonephritis, cirrhosis, and scleroderma, particularly
CC proliferative disorders in which loss of a TGF-beta autocrine or
CC paracrine signaling is implicated. They can also be used to
CC modulate spermatogenesis and to treat disorders of the female
CC reproductive system which lead to infertility including polycystic
CC ovarian syndrome. The products can also be used for detection,
CC diagnosis and drug screening.
XX
SQ Sequence 421 AA;

Query Match 37.1%; Score 508; DB 20; Length 421;
Best Local Similarity 42.5%; Pred. No. 1.8e-43;
Matches 114; Conservative 50; Mismatches 78; Indels 26; Gaps 10;

```
Qy       1 DVLEDSETWDQATGTKTFLVSQDI--RDEGWETLEVSSAVKRWVRADSTTNKNKLEVTVQ 58
           :||| ::  ::   :   ||| :|   :  ||| :|: | :|| ::  :|  ::||: ::
Db     169 EVLESADGSEEER-SMLVLVSTEIYGTNSEWETFDVTDATRRWQKSGPST--HQLEIHIE 225

Qy      59 SHRESCD-----TLDISVPPGSKNLPFFVVFSNDRSNGTKETRLDLLKEMIGHEQETMLV 113
           | :   :      |:| :   :|: |  ||||:|:||  || : : | |:| |||:  |
Db     226 SRQNQAEDTGRGQLEIDMSAQNKHDPLLVVFSDDQSN-DKEQK-EELNELITHEQDLDLD 283

Qy     114 KTAKNAYQGAGE----SQEEEGLDGYTAVGPLLARRKRSTGASSHCQKTSLRVNFEDIGW 169
             |   : |  |          :|  :|       | |     ::|:|| | ::|::|||
Db     284 SDA--FFSGPDEEALLQMRSNMIDDSSA-------RIRRNAKGNYCKKTPLYIDFKEIGW 334

Qy     170 DSWIIAPKEYDAYECKGGCFFPLADDVTPTKHAIVQTLVHLKFPTKVGKACCVPTKLSPI 229
           |||||||  |:||||:| | :|||: :|||||||:| |||||   |  |||||||| ||
Db     335 DSWIIAPPGYEAYECRGVCNYPLAEHLTPTKHAIIQALVHLKNSQKASKACCVPTKLDPI 394

Qy     230 SILYKDDMGVPTLKYHYEGMSVAECGCR 257
```

```
                ||||  | || | |: ||||:|:|||||
Db     395 SILYL-DKGVVTYKFKYEGMAVSECGCR 421
```

Search completed: April 2, 2003, 14:49:34
Job time : 77 secs